

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:57:59 ; Search time 75 Seconds
(without alignments)
2140.173 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPLDRR.....PLEVITFEFAVGAKQAAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4075	100.0	800	14	US-10-051-909-32
2	3774	92.6	745	12	US-10-310-154-718
3	3517	86.3	747	14	US-10-051-902-2
4	3517	86.3	747	14	US-10-051-909-2
5	2834	69.5	740	14	US-10-051-909-37
6	2674	65.6	737	14	US-10-051-902-8
7	2674	65.6	737	14	US-10-051-909-8
8	2584	63.4	729	14	US-10-051-902-29
9	2584	63.4	729	14	US-10-051-909-29
10	1670.5	41.0	486	14	US-10-051-902-10
11	1670.5	41.0	486	14	US-10-051-909-10
12	1424	34.9	345	14	US-10-051-902-14
13	1424	34.9	345	14	US-10-051-909-14
14	943	23.1	228	14	US-10-051-902-16
15	943	23.1	228	14	US-10-051-909-16

16	587.5	14.4	457	12	US-10-369-493-23324	Sequence 23324, A
17	562	13.8	131	14	US-10-051-902-6	Sequence 6, Appli
18	562	13.8	131	14	US-10-051-909-6	Sequence 6, Appli
19	554.5	13.6	487	9	US-09-795-693-27	Sequence 27, Appl
20	554.5	13.6	487	12	US-10-391-399-45	Sequence 45, Appl
21	554.5	13.6	487	14	US-10-095-139-14	Sequence 14, Appl
22	554.5	13.6	487	15	US-10-156-239-27	Sequence 27, Appl
23	554.5	13.6	487	15	US-10-162-102-46	Sequence 46, Appl
24	548	13.4	488	12	US-10-162-012-46	Sequence 4, Appli
25	548	13.4	488	14	US-10-094-059-4	Sequence 5, Appli
26	548	13.4	488	15	US-10-170-528-5	Sequence 46, Appl
27	548	13.4	488	15	US-10-062-960B-4	Sequence 4, Appli
28	548	13.4	488	15	US-10-062-960B-4	Sequence 4, Appli
29	548	13.4	488	15	US-10-144-624-4	Sequence 7, Appli
30	547	13.4	486	9	US-09-860-232A-7	Sequence 722, App
31	534	13.1	580	12	US-10-310-154-722	Sequence 23371, A
32	533	13.1	461	12	US-10-369-493-23371	Sequence 20, Appl
33	531.5	13.0	535	15	US-10-156-239-20	Sequence 20, Appl
34	531.5	13.0	535	15	US-10-199-485-20	Sequence 20, Appl
35	531.5	13.0	535	15	US-10-199-485-20	Sequence 24, Appl
36	505	12.4	473	12	US-10-369-493-23097	Sequence 20, Appl
37	496	12.2	523	14	US-10-051-902-24	Sequence 24, Appl
38	496	12.2	523	14	US-10-051-909-24	Sequence 20, Appl
39	489.5	12.0	513	14	US-10-051-902-20	Sequence 20, Appl
40	489.5	12.0	513	14	US-10-051-909-20	Sequence 28, Appl
41	489	12.0	529	14	US-10-051-902-28	Sequence 28, Appl
42	489	12.0	529	14	US-10-051-909-28	Sequence 40, Appl
43	488	12.0	548	11	US-09-774-381-40	Sequence 30, Appl
44	488	12.0	549	14	US-10-051-902-30	Sequence 30, Appl
45	488	12.0	549	14	US-10-051-909-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-10-051-909-32
; Sequence 32, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-909-32

Query Match	100.0%;	Score 4075;	DB 14;	Length 800;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 800;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	IRSGSWLAVQTPFTPLDRRLPSVVLALPGPLPPASCSQEPVTSDDLKMGSCAV	60	
Db	1	IRSGSWLAVQTPFTPLDRRLPSVVLALPGPLPPASCSQEPVTSDDLKMGSCAV	60	
QY	61	LVAIVASIGNLQGNDAITAAVLYIKKEFQONEPTVEGLIVSMISLIGATITTFSGP	120	
Db	61	LVAIVASIGNLQGNDAITAAVLYIKKEFQONEPTVEGLIVSMISLIGATITTFSGP	120	
QY	121	LSDSIGRRPMLILSSILYFSGLIMLWSPNVYLLARFVDGFGIGLAVTVLYISEIA	180	

Db 121 LSDSIGRRPMLILSSILYFSSGLIMLMSPNVYVLLARFVDFGIGLAATLVPLVISEIA 180
QY 181 PSFIRGLNLTLPQFSSGGMFLSYCNVFGMSLSPSPDWIRIMLGVLAIPLSLFFPGLTIFYL 240
Db 181 PSFIRGLNLTLPQFSSGGMFLSYCNVFGMSLSPSPDWIRIMLGVLAIPLSLFFPGLTIFYL 240
QY 241 PESPRWLVSCKGRMAEAKVLOKLRGDDVSGELSLLEGLVGGDTSEIYIIGPATEAA 300
Db 241 PESPRWLVSCKGRMAEAKVLOKLRGDDVSGELSLLEGLVGGDTSEIYIIGPATEAA 300
QY 301 DDLVTDGDKQITLYGPEEGQSWIARPSKPIMLGSLSLASRHGSMWNQSVPLMDPIVT 360
Db 301 DDLVTDGDKQITLYGPEEGQSWIARPSKPIMLGSLSLASRHGSMWNQSVPLMDPIVT 360
QY 361 LFGSVHNMPOAGGSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEEYASDAGGD 420
Db 361 LFGSVHNMPOAGGSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEEYASDAGGD 420
QY 421 YEDNLHSPLLSRQATGAEGKDIYHGHGSAALSMRQOTLLEGGDVSSSTDIGGQQLAM 480
Db 421 YEDNLHSPLLSRQATGAEGKDIYHGHGSAALSMRQOTLLEGGDVSSSTDIGGQQLAM 480
QY 481 KWSKKEGNGRKEGGFKRVYLHOGVPGSRGSIIVSLPGGQVFESEFVHAALVSQSA 540
Db 481 KWSKKEGNGRKEGGFKRVYLHOGVPGSRGSIIVSLPGGQVFESEFVHAALVSQSA 540
QY 541 LFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRRALVGVGIQILOQFAGINGVL 600
Db 541 LFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRRALVGVGIQILOQFAGINGVL 600
QY 601 YVTPQILEQAGVAVILSKFGLSSASISLISLTTLLMPCIGFAMLLMDLSGRFFLLG 660
Db 601 YVTPQILEQAGVAVILSKFGLSSASISLISLTTLLMPCIGFAMLLMDLSGRFFLLG 660
QY 661 TIFILIASLVLVNSNLDLGTALHALSTVSIVVYFCCFVWGFGPIPNILCAEIPFRV 720
Db 661 TIFILIASLVLVNSNLDLGTALHALSTVSIVVYFCCFVWGFGPIPNILCAEIPFRV 720
QY 721 RGLCTAICAFTEWIGDIIVTYSLPVWMLNAGLAGVFSIYAVCLISFVFLKVPETKGM 780
Db 721 RGLCTAICAFTEWIGDIIVTYSLPVWMLNAGLAGVFSIYAVCLISFVFLKVPETKGM 780
QY 781 PLEVITTEFFAVGAKQAAAKA 800
Db 781 PLEVITTEFFAVGAKQAAAKA 800

RESULT 2
US-10-310-154-718

; Sequence 718, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Beil, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shishieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.

; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanqun
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 718
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Zea mays
US-10-310-154-718

10348712d

Query Match 92.6%; Score 3774; DB 12; Length 745;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 742; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 56 MSGAVLVAIVASIGNLQGMNDATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAVLVAIVASIGNLQGMNDATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFSSGLIMLMSPNVYVLLARFVDFGIGLAATLVPLV 175
Db 61 TFSGPLSDSIGRRPMLILSSILYFSSGLIMLMSPNVYVLLARFVDFGIGLAATLVPLV 120
QY 176 ISEIAPSEIRGLNLTLPQFSSGGMFLSYCNVFGMSLSPSPDWIRIMLGVLAIPLSLFFPG 235
Db 121 ISEIAPSEIRGLNLTLPQFSSGGMFLSYCNVFGMSLSPSPDWIRIMLGVLAIPLSLFFPG 180
QY 236 TIFYPESPRLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLVGGDTSEIYIIGP 295
Db 181 TIFYPESPRLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLVGGDTSEIYIIGP 240
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKPIMLGSLSLASRHGSMWNQSVPLM 355
Db 241 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKPIMLGSLSLASRHGSMWNQSVPLM 300
QY 356 DPIVTLFGSVHNMPOAGGSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEEYASD 415
Db 301 DPIVTLFGSVHNMPOAGGSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEEYASD 360
QY 416 GAGGDYEDNLHSPLLSRQATGAEGKDIYHGHGSAALSMRQOTLLEGGDVSSSTDIGGG 475
Db 361 GAGGDYEDNLHSPLLSRQATGAEGKDIYHGHGSAALSMRQOTLLEGGDVSSSTDIGGG 420
QY 476 WQIAWKSEKEGNGRKEGGFKRVYLHOGVPGSRGSIIVSLPGGQVFESEFVHAAL 535
Db 421 WQIAWKSEKEGNGRKEGGFKRVYLHOGVPGSRGSIIVSLPGGQVFESEFVHAAL 480

QY 536 VSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQOFAG 595
Db 481 VSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQOFAG 540
QY 596 INGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSGR 655
Db 541 INGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSGR 600
QY 656 FLLGTIPILIASLVILVSNLIDGLTALHALLSTVSIVYFCCFVMGFGPIPNILCAEI 715
Db 601 FLLGTIPILIASLVILVSNLIDGLTALHALLSTVSIVYFCCFVMGFGPIPNILCAEI 660
QY 716 FPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFVFLKVP 775
Db 661 FPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFVFLKVP 720
QY 776 ETKGMPLEVITEFFAVGAKQAATA 800
Db 721 ETKGMPLEVITEFFAVGAKQAATA 745

RESULT 3
US-10-051-902-2
; Sequence 2, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)
; NAME/KEY: UNSURE
; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
; US-10-051-902-2

Handwritten signature/initials

Query Match 86.3%; Score 3517; DB 14; Length 747;
Best Local Similarity 91.6%; Pred. No. 1.7e-305;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY 56 MSGAVLVAIVASIGNLLOGWMDNATIAAVLYIKKEFOLNEPTVEGLIVSMISLIGATVIT 115
Db 1 MCGAVVAIAIASIGNLLOGWMDNATIAAVLYIKKEFOLNEPTVEGLIVSMISLIGATVIT 60
QY 116 TFGPLSDSIGRRPMLILSSILYFSGILMLSPNVYVLLARFVDGFGIGLAVTVLPLV 175
Db 61 TSPGPRADCVGRPRMLVASAVLYFVSGVLMVLPVIVYLLARLIDGFGIGLAVTVLPLV 120

QY 176 ISEIAPSEIR-GLINTLPQFSG-SGGMFLSYCVFSGMSLSPSPDWRLMLGVALPSLFEF 233
Db 121 ISETAPHRXSWGXNLTLPQFIVGXGGMFLSYCVFSGMSLSPSPDWRLMLGVALPSLXYF 180
QY 234 GLTIFYLPSPRWLVSKGRMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIEYII 293
Db 181 GLTIFYLPSPRWLVSKGRMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIEYII 240
QY 294 GPATEAADLVTDDKEQITLYGPEEGSWIARPSKPIMLGSLSLASRHSVMNQSV 353
Db 241 GPATEAADLVTDDKEQITLYGPEEGSWIARPSKPIMLGSLSLASRHSVMNQSV 300
QY 354 LMDPIVTLFGSVHNNPQAGGSMRSTLFPNFGSMFSVTDOHAKNEQWDEENLHRDDEYA 413
Db 301 LMDPIVTLFGSVHNNPQAGGSMRSTLFPNFGSMFSVTDOHAKNEQWDEENLHRDDEYA 360
QY 414 SDGAGDYEDNLHSPILSRQATGAEGKDIVHGHRSALSMRRQTLLEGGDVSSTDIG 473
Db 361 SDGAGDYEDNLHSPILSRQATGAEGKDIVHGHRSALSMRRQTLLEGGDVSSTDIG 420
QY 474 GGMQLAMKWESEKENGKKEGFKRVYVTHQEGVDSRRGSIVSLPGGADVSESEFVHAA 533
Db 421 GGMQLAMKWESEKENGKKEGFKRVYVTHQEGVDSRRGSIVSLPGGADVSESEFVHAA 480
QY 534 ALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQOF 593
Db 481 ALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQOF 540
QY 594 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSG 653
Db 541 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSG 600
QY 654 RFLGLTIPILIASLVILVSNLIDGLTALHALLSTVSIVYFCCFVMGFGPIPNILCA 713
Db 601 RFLGLTIPILIASLVILVSNLIDGLTALHALLSTVSIVYFCCFVMGFGPIPNILCA 660
QY 714 EIPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFVFLK 773
Db 661 EIPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFVFLK 720
QY 774 VPETKMPLEVITEFFAVGAKQAATA 800
Db 721 VPETKMPLEVITEFFAVGAKQAATA 747

Handwritten signature/initials

RESULT 4
US-10-051-909-2
; Sequence 2, Application US/10051909
; Publication No. US2002019217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)
; NAME/KEY: UNSURE

```

; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
; US-10-051-909-2
```

Query Match 86.3%; Score 3517; DB 14; length 747;
Best Local Similarity 91.6%; Pred. No. 1.7e-305;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

```

QY 56 MSGAVLVAIVASIGNLLOGWDNATIAAVALYIKKEFOLONEPTVEGLIVMSLIGATIVT 115
Db 1 MGAVMVAIAASIGNLLOGWDNATIAGAVLYIKKEFNLOSEPLIEGLIVAMFLIGATIVT 60

QY 116 TFSGPLSDSIGRRPMLILSSILYFSSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TSPGRADCVGRPMLVASAVLYFVSGLWMLAPIVYILLARLIDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIR-GLLNTLPQFSG-SGGMFLSYCMVFGMSLSPSPDWIRIMLGVALIPSLFFP 233
Db 121 ISETAPHRKXSWGXXNTLPQFIGYAGMFLSYCMVFGMSLMPKPDWRIMLGVALISPLXYF 180

QY 234 GLTIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIYYII 293
Db 181 GLTIFYLPESPRWLVSKGMAEAKVQRLRGREDVSXEXALLVEGLGVGKDIRIXEYII 240

QY 294 GPAITEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVP 353
Db 241 GPAITEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVP 300

QY 354 LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDEEYA 413
Db 301 LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDEEYA 360

QY 414 SDGAGDYEDNLHSPILSRQATGABKDIVHHGRGALSRRQTLLEGGDVSSSTDIG 473
Db 361 SDGAGDYEDNLHSPILSRQATGABKDIVHHGRGALSRRQSLLEGGDVSSSTDIG 420

QY 474 GGWQLAMKWKSEKEGENGRKEGFRVYLIHQEGVPGSRGSIYSLPGGADVPEGSEFVHAA 533
Db 421 GGWQLAMKWKSEKEGENGRKEGFRVYLIHQEGVPGSRGSIYSLPGGADVPEGSEFVHAA 480

QY 534 ALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIOILOQF 593
Db 481 ALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIOILOQF 540

QY 594 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSG 653
Db 541 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSG 600

QY 654 RRFLLGTIPILIASLVILVSNLIDGLTAHALSTVSVIYVFCCFWGMFGPIPNILCA 713
Db 601 RRFLLGTIPILIASLVILVSNLIDGLTAHALSTVSVIYVFCCFWGMFGPIPNILCA 660

QY 714 EIFPTRVRGLCIAICATFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLK 773
Db 661 EIFPTRVRGLCIAICATFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLK 720

QY 774 VPETKMPLEVITEFFAVGAKQAAKA 800
Db 721 VPETKMPLEVITEFFAVGAKQAAKA 747
```

RESULT 5
US-10-051-909-37
; Sequence 37, Application US/10051909

```

; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR APPLICATION NUMBER: 2002-01-17
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-051-909-37
```

Query Match 69.5%; Score 2834; DB 14; length 740;
Best Local Similarity 74.5%; Pred. No. 2.2e-244;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

```

QY 56 MSGAVLVAIVASIGNLLOGWDNATIAAVALYIKKEFOLONEPTVEGLIVMSLIGATIVT 115
Db 1 MAGAVLVAIAASIGNLLOGWDNATIAGAVLYIKKEFNLOSEPLIEGLIVAMSLIGATIT 60

QY 116 TFSGPLSDSIGRRPMLILSSILYFSSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFSGAVADSFGRRPMLIASAVLYFVSGLWMLAPNVYVLLARLIDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALIPSLFFGL 235
Db 121 ISETAPDIRGLNTLPQFSGSGGMFLSYCMVFGMSLMPQPDWRIMLGVALISPLIYFAL 180

QY 236 TIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIYYIIGP 295
Db 181 TIFYLPESPRWLVSKGMAEAKVLOGLRGREDVSGEMALLVEGLGVGKDIRIXEYIIGP 240

QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVPLM 355
Db 241 DDELADDEGLAP-DPEKIXLYGPEEGLSWVARPVHQQSALGSALGLISRHGSMVNSQKPLV 299

QY 356 DPVITLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDEEYASD 415
Db 300 DPVITLFGSVHENMPQAGSMRSTLFPNFGSMFVSAEQQAKGDWDAES-QREGEDYGSD 358

QY 416 GAGGDYEDNLHSPILSRQATGABKDI-VHHGRGALSRRQTLLEGGDVSSSTDIG 474
Db 359 HGGDDIEDLSQPLISROATSVEKEIAAPHGSIMGAVG--RSSSLMQGGEAVSSMGIGG 416

QY 475 GWQLAMKWKSEKEGENGRKEGFRVYLIHQEGVPGSRGSIYSLPGGADVPEGSEFVHAAA 534
Db 417 GWQLAMKWKTEREGADKEGEGFORIYLHEEGVTGDRGSIYSLP-GGDVPPGGEFVQAAA 475

QY 535 LVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIOILOQFA 594
Db 476 LVSQPALYSKELMEQBLAGPAMVHPQAVAKGPXWADLFEPGVKHALFVGIGIOILOQFA 535

QY 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSGR 654
Db 536 GINGVLYTTPQILEQAGVLLANIGLSSSSASILISGLTLLMLPSIGIMRMLDMSGR 595

QY 655 RFLLLGTIPILIASLVILVSNLIDGLTAHALSTVSVIYVFCCFWGMFGPIPNILCAE 714
Db 596 RFLLLATIPILIALAILLVNILDVGTMTASISTVSVIYVFCFVWGMFGPIPNILCAE 655

QY 715 IFPTRVRGLCIAICATFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLKV 774
Db 656 IFPTVRGICIAICATFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLIAFLVFMKV 715
```

QY 775 PETKMPLEVITEFFAVGAKQA 796
Db 716 PETKMPLEVITEFFSVGAKQA 737

RESULT 6
US-10-051-902-8

; Sequence 8, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR APPLICATION NUMBER: 2002-01-17
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-051-902-8

Query Match 65.6%; Score 2674; DB 14; Length 737;
Best Local Similarity 69.9%; Pred. No. 4.7e-230;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASIGNLQGMWDNATIAAALYIKKEFQONEPTVEGLIVSMGLIGATIVT 115
Db 1 MKGAVLVAIAASIGNFLQGMWDNATAGANGYIKKDLALGT--TMERLAVGMSLIGATIVT 58
QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 59 TCSGPIADWLGRPRMIISSVLYFLGLGLVLMWSPNVYVLCIARLLDGFGLAVTLVPLY 118
QY 176 ISEIAPSEIRGLINTLPQFSSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALPISLFFEGL 235
Db 119 ISETAPSEIRGLINTLPQFSSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALPISLFFEGL 178
QY 236 TIFYLPESPRWLVSCKRMAEAKKVLQKRGKDDVSGELSLLEGLVEGGDTSEIYYIGP 295
Db 179 TIFYLPESPRWLVSCKRMAEAKKVLQKRGKDDVSGELSLLEGLVEGGDTSEIYYIGP 238
QY 296 ATEAADLVTGDKEQITLYGPEEGQSWIARPSKPIMLGSVLSLARHGSMTVNQSVPLM 355
Db 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPVGTQSSIG---LASHHGSIIINQSMPLM 294
QY 356 DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEEYA 413
Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLFPNFGSMFSTAEPAKIEQWDEESLQREREDYM 354
QY 414 SDGAGDYEEDNLHSPILSRQATGAEGKDIYHHGHSAL-SMRQTLGEG-GDGVSTSD 471
Db 355 SDATRGSDSDNLHSPILSRQATGAEGKDIYHHGHSAL-SMRQTLGEG-GDGVSTSD 413
QY 472 IGGWQOLAMKWEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGADVFESEFVH 531
Db 414 IGGWQOLAMKWTDK-GEDEKQGGFKRIYLYHEEGVSASRRGSIIVSIPEG------EFVQ 466
QY 532 AAALVSQSALFSKGLAEPKMSDAAVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQ 591
Db 467 AAALVSQPALYSKELIDHPVGPAMVHPSETASKGSPWKALLFEPGVKHALVGVGIQILQ 526
QY 592 QFAGINGVLYTTPQILEQAGVAIVLSKGLSSASASILISLTLMLPCIGFAMLMIDL 651
Db 527 QFSGINGVLYTTPQILEQAGVAIVLSKGLSSASASILISLTLMLPCIGFAMLMIDL 586
QY 652 SGRFLLLTGITPILIASLIVVSNLIDGLTAAHALSTVSVIVYFCFVWVGFPINIL 711

Db 587 SGRQOLLTLTIPVLIVSLIITLVIGSLVNFGNVAHAISTVCVVYFCCFWMGXGPIPNIL 646
QY 712 CAEIFPTRVRCICIAICFTFWIGDITVYSLPVMNAILGAVPSIYAVVCLISFVFEV 771
Db 647 CSEIFPTRVRCICIAICALVFWIGDIIITYSLPVMGLSLGLGVFAIYAVVCFISWIFV 706
QY 772 LKVPETKMPLEVITEFFAVGAKQAAA 798
Db 707 LKVPETKMPLEVITEFFSVGAKQAAS 733

RESULT 7
US-10-051-909-8

; Sequence 8, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR APPLICATION NUMBER: 2002-01-17
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-051-909-8

Query Match 65.6%; Score 2674; DB 14; Length 737;
Best Local Similarity 69.9%; Pred. No. 4.7e-230;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASIGNLQGMWDNATIAAALYIKKEFQONEPTVEGLIVSMGLIGATIVT 115
Db 1 MKGAVLVAIAASIGNFLQGMWDNATAGANGYIKKDLALGT--TMERLAVGMSLIGATIVT 58
QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 59 TCSGPIADWLGRPRMIISSVLYFLGLGLVLMWSPNVYVLCIARLLDGFGLAVTLVPLY 118
QY 176 ISEIAPSEIRGLINTLPQFSSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALPISLFFEGL 235
Db 119 ISETAPSEIRGLINTLPQFSSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALPISLFFEGL 178
QY 236 ATEAADLVTGDKEQITLYGPEEGQSWIARPSKPIMLGSVLSLARHGSMTVNQSVPLM 355
Db 179 TIFYLPESPRWLVSCKRMAEAKKVLQKRGKDDVSGELSLLEGLVEGGDTSEIYYIGP 238
QY 296 DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEEYA 413
Db 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPVGTQSSIG---LASHHGSIIINQSMPLM 294
QY 414 SDGAGDYEEDNLHSPILSRQATGAEGKDIYHHGHSAL-SMRQTLGEG-GDGVSTSD 471
Db 355 SDATRGSDSDNLHSPILSRQATGAEGKDIYHHGHSAL-SMRQTLGEG-GDGVSTSD 413
QY 472 IGGWQOLAMKWEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGADVFESEFVH 531
Db 414 IGGWQOLAMKWTDK-GEDEKQGGFKRIYLYHEEGVSASRRGSIIVSIPEG------EFVQ 466
QY 532 AAALVSQSALFSKGLAEPKMSDAAVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQ 591

Db 467 AAALVSQPALYSKELIDGHPVGPAMVHPSETASKG;SWKALIEPBGKHALVGVGIQLIQ 526
QY 592 QPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLMDL 651
Db 527 QFSGINGVLYTTPQILEEAGVEVLSDIGIGSESASFLISAFITFLMLPCIGVAMKMDV 586
QY 652 SGRFLLGTIPILIASLVILVSNLIDGLAHALSTVSVIVYFCCFVMGFGPIPNIL 711
Db 587 SGRQOLLTTIPVLIVSLIILVIGSLVFNFGVNAHAALSTVCVVYFCFVMGYGPIPNIL 646
QY 712 CABIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMINAIGLAVFSIYAVVCLISFVYF 771
Db 647 CSEIFPTRVRGLCIAICALVFWIGDIIITYSLPVMIGSLGCVFAIYAVVCFISWIFVF 706
QY 772 LKVPETKGMPLLEVITEFFAVGAKQAAA 798
Db 707 LKVPETKGMPLLEVISEFFSVGAKQAAAS 733

RESULT 8

US-10-051-902-29
; Sequence 29, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-902-29

Query Match 63.4%; Score 2584; DB 14; Length 729;
Best Local Similarity 68.8%; Pred. No. 5.2e-222;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
QY 56 MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQONEPTVEGLIVSMISLIGATIVT 115
Db 1 MSGAVLVAIAAAGNLLQGMNATIAGAVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TCSGCVADWLGRPMILSSILYFVGSVLMLWSPNVYVLLGRLLDGFVGVLVTLVPIY 120
QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRIMGLVLAIPSLFFFG 235
Db 121 ISETAPEIRGLNLTLPQFTGSGGMFLSYCMVFGMSLSPSPDWIRIMGLVLAIPSLVFFFL 180
QY 236 TIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIEYIIGP 295
Db 181 TVFFLPESPRWLVSKGMAEAKVLOKLRGREDVSGEMALLVEGLIGGETTIEEYIIGP 240
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGS-MVNQSVPL 354
Db 241 ADEVTDHDIADVDDQIKLYGAEEGLSWARPVKG----GSTMSVLSRHGSTMSRRQSL 296
QY 355 MDPIVTLFGSVHNMPOAGSGMRSTLFPNFGSMFSVTDOHAKNEQWDEENLHRDDEEYAS 414
Db 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPFHGSMFSVCGNQPRHEDWDEENLVGEGEDYPS 355
QY 415 DGAGGDYEDNLHSPILSRQATGAEGKDIVHHGRGSALSMKRQTLLEGEGDGVSTDIGG 474

Db 356 D-HGDDSEDDLHSPILSRQTSME-KDMPTHAGTLSTFRHGSQVQAGQEBGAGSMGIGG 413
QY 475 GWOLAMKWESEKEGENRKEGGFKRVYLHOEGVPGSRGSIIVSLPGGADVFESEFVHAAA 534
Db 414 GWQVAMKWTREDESGQKE-----EGFPGSRGSIIVSLPGGDTGE-ADFVQASA 462
QY 535 LVSQALFSKGLAEPRMSDAMVHPSEVAACKSRWKDLFEPGVRRALLVGVGIQLQQA 594
Db 463 LVSQPALYSKDLKEHTIGPAMVHPSE-TTKGSIMWDLHDPGVKRALVGVGIQLQOFS 521
QY 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYTTPQILEQAGVGILLSNMGISSSASLISALTTFVMLPAIYAVARLMDLSGR 581
QY 655 RFLLTGTIPILIASLVILVSNLIDGLAHALLSTVSVIVYFCCFVMGFGPIPNILCAE 714
Db 582 RLLLTTPILIASLVILVSNLVHMSIVHAVLSTVSVLYFCFVMGFGPAPNILCSE 641
QY 715 IFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMINAIGLAVFSIYAVVCLISFVYFELKV 774
Db 642 IFPTRVRGLCIAICALTFWICDIIVTYSLPVLKSLIGLAVFGMVAIVCCISWVVFPIKV 701
QY 775 PETKGMPLLEVITEFFAVGAKQAAA 798
Db 702 PETKGMPLLEVITEFFSVGARQAAA 725

RESULT 9

US-10-051-909-29
; Sequence 29, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helencjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-909-29

Query Match 63.4%; Score 2584; DB 14; Length 729;
Best Local Similarity 68.8%; Pred. No. 5.2e-222;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
QY 56 MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQONEPTVEGLIVSMISLIGATIVT 115
Db 1 MSGAVLVAIAAAGNLLQGMNATIAGAVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TCSGCVADWLGRPMILSSILYFVGSVLMLWSPNVYVLLGRLLDGFVGVLVTLVPIY 120
QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRIMGLVLAIPSLFFFG 235
Db 121 ISETAPEIRGLNLTLPQFTGSGGMFLSYCMVFGMSLSPSPDWIRIMGLVLAIPSLVFFFL 180
QY 236 TIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIEYIIGP 295
Db 181 TVFFLPESPRWLVSKGMAEAKVLOKLRGREDVSGEMALLVEGLIGGETTIEEYIIGP 240
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGS-MVNQSVPL 354

Db	241	ADEVTDHDIANDVKDQIKLYGAEEGLSWAVRPVKG----	GSTMSVLSRHGSTMRRQGS	296
QY	355	MDPIVTLFGSVHENMPQAGGSMRSTLEPNFGSMFSVT	DQAKNEOWDEENIHRDEEYAS	414
Db	297	IDPLVTLFGSVHEKMPDT-GSMRSALFBHFGSMFSV	GNQPRHEDWDEENLVGEDGYPS	355
QY	415	DGAGGDYEDNLHSPILSRQATGAEGKDIVHHGRG	SALSMRRQTLLEGGDVSTDIG	474
Db	356	D-HGDDSEDDLHSPILSRQTSME-KDMHTAHGTL	STFRHGSQVQGAQGEAGSMGIG	413
QY	475	GMQLAMKMKSEKEGENGRKEGGFKRVYLHQEGV	PGSRGSI VSLPGGADVFESEFVHAA	534
Db	414	GMQVAMKMTTEREDESQKE-----EGFP	SGRSISVSLPGGDGTGE-ADFVQASA	462
QY	535	LVSQSALFSKGLAEPRMSDAAWHPSEVAAKGSR	KDLEPGVRALLVGVIQILQOFA	594
Db	463	LVSQPALYSKDLKEHTIGPAMWHPSE-TTKGSI	WHDLHPGVKRALVVGVIQILQOFS	521
QY	595	GINGVLYTTPQILEQAGVAVILSKFGISSASASIL	ISSLTTLMLPCI GFAMLLMDLSGR	654
Db	522	GINGVLYTTPQILEQAGVILLSNMGISSSSAS	LLISALTTFVMLPAIAVAKRIMDLSCR	581
QY	655	RFLLLGTIPILIASLVILVNSNLIDGLTAHAL	LSTVSVIVYFC CFVMGFGPIPNILCAE	714
Db	582	RTLLLTIPILIASLLVLVINSLVHMSNIVHA	VLSTVSVLVYFCFVMGFGPAPNILLSE	641
QY	715	IFPTRVRGLCIAICAFTFWIGDIIVTYSLEP	MLNAILAGLAVFSIYAVVCLISFVFFLVK	774
Db	642	IFPTRVRGICIAICALTFWICDIIVTYSLEP	LLKSIGLAVFGMYAIVCCISWVFFIKV	701
QY	775	PETKGMPLLEVITTEFPVAGAKQAAA	798	
Db	702	PETKGMPLLEVITTEFPVSGARQAEA	725	

```

RESULT 10
US-10-051-902-10
; Sequence 10, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-902-10

```

Query Match	41.0%;	Score 1670.5;	DB 14;	Length 486;
Best Local Similarity	68.1%;	Pred. No. 1.6e-140;		
Matches 333;	Conservative 57;	Mismatches 84;	Indels 15;	Gaps 8;

[illegible][illegible]

```

RESULT 11
US-10-051-909-10
; Sequence 10, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-909-10

```

Query Match	41.0%;	Score 1670.5;	DB 14;	Length 486;
Best Local Similarity	68.1%;	Pred. No. 1.6e-140;		
Matches 333; Conservative	57;	Mismatches 84;	Indels 15;	Gaps 8;

QY	308	DKEOITLLYPBEGQSWIARPSKPIMLGSLSLASRHGSMVNQSVPLMDPIVTLFGSVHE	367
		: : : : : : : : : : : :	
Db	5	EKDQIKLYPBGQGSWVARPVAGPNSVG---LVSRKGSMANPS-SLVDPVLTLFGSVHE	59
QY	368	NMEQAGGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDDEEYASDAGGDYEDNLHS	427
		: : : : : : :	
Db	60	KLPETG----STLFPHFSGSMFSVGNOPRNEDWDEESLAREGDDYSD--AGSDSDNLQS	113
QY	428	PLLSRQATGAEGKDIVHHGHRGASALSMRQTL--GEGGDVSSTDIGGQOLAWKMSKE	486
		: : : : : :	
Db	114	PLISRQTTSLD-KDIPPHAHSNLA-SMRQGSLLHGNSGEPGTGIGGQOLAWKMSERE	171
QY	487	GENSRKEGGEFKRYVLLHQEGVPGSRKGSIVSLPGGGDVPEGSEFVAHAALVSOSALFSKGL	546
		: : : : : : :	
Db	172	GPDGKEGGEFKRIYLLHQDGSGSRKGSVSVLP-GGDLPTDSEVVQAALVSQPALYNEDL	230
QY	547	AEPNMSDAAVHPSEVAAKGSRWKDLFEPGVRRLVLVGIOILOQFAGINGVLYTPOI	606
		:	
Db	231	MROFPVGPAMIHPSSETIAKGPWSWDLFEEGVKHALIVGVMQILOQFSGINGVLYTPOI	290

QY	607	LEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSGRFFLLGTPIILI	666
		: : : : : : : : : : :	
Db	291	LEQAGVGYLLSSLGSTSSSFLISAVTTLMLPCIAIAMRLMDISGRRTLLSTIPVLI	350
QY	667	ASVILVVSNI,IDLGLAHALLSTVSIVVFCEVMGFGPIPNILCAEIPFTRVGLCIA	726
		: : : : : : : : : : : :	
Db	351	AALLILVGLSLVDLSTANASISTISVIVFCEVMGFGPIPNILCAEIPFTRVGLCIA	410
QY	727	ICAFETWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVFLKVPETKMBLEVIT	786
		: : : : : : :	
Db	411	ICALLFWICDIIVTYTLPVMLNSVGLAGVFGIYAVVCFIAWVFVFLKVPETKMBLEVII	470
QY	787	EFFAVGAKQ 795	
		:	
Db	471	EFFSVGAKQ 479	

RESULT 12

```

US-10-051-902-14
; Sequence 14, Application US/10051902
; Publication No: US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-902-14

```

Query Match	34.9%;	Score 1424;	DB 14;	Length 345;
Best Local Similarity	82.9%;	Pred. No. 1.1e-118;		
Matches 281;	Conservative 24;	Mismatches 32;	Indels 2;	Gaps 2
QY 462	EGDGVSTDIGGQWLAWKWKSEKEGNGRKEGGFIRVYLHQEGVPGSRGSIIVSLPGGG	521		
DB 4	EGGEAVSSTGIGGGQWLAWKWSERQGEDGKEGGFIRIYLYHQEGVADSRGRGSIVSLPGGG	63		
QY 522	DVEE-GSEFVHAALVQSALFSKGLAEPRM-SDAAMVHPSEVAAGSRWKDLFEPGVR	579		
DB 64	DATQGGSGFIHAALVSHSALYSKDLMEERMAAGPAMIHPLAAPKGSITWKDLFEPGVR	123		
QY 580	ALLVGVGIQILQQFAGINGVLYTTPQILQAGVAVILSKFGLSSASASTLISSTLTLLML	639		
DB 124	ALFVGVIQMLQQFAGINGVLYTTPQILQAGVAVILSNLGLSSASASTLISSTLTLLML	183		
QY 640	PCIGFAMLLMDLSGRFLLGTIPLILASVLIVVSNLIDLGLTAHALSTVSVIVYFCC	699		
DB 184	PSIGVAMRLMDISGRFLLGTIPLILASVLIVGVNVINLSTVPHAVLSTVSVIVYFCC	243		
QY 700	FVMGFGPIPNILCAEIFPTRVRGLCTAICAFTFWIGDIIVTYSLPVNLAIAGLGVFSIY	759		
DB 244	FVMGFGPIPNILCAEIFPTRVRGVCIAICALTFWICDIIVTYSLPVNLAIAGLGVFGIY	303		
QY 760	AVVCLISFVFVLKVPETKGMPLVITEFFAVGAKCAAA	798		
DB 304	AVVCCIAFVFVYLKVPETKGMPLVITEFFAVGAKCAOA	342		

RESULT 13

US-10-051-909-14
; Sequence 14, Application US/10051909

```

; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B0163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-14

```

[illegible]

RESULT 14

```

US-10-051-902-16
; Sequence 16, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-902-16

```


Query Match 23.1%; Score 943; DB 14; Length 228;
Best Local Similarity 80.6%; Pred. No. 6.6e-76;
Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 574 EPGVRRALLVGVGIQLIQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISL 633
Db 1 EPGVKHALFVIGLQILQFAGINGVLYTTPQILEQAGVAVILSNIGLSSSSASILISAL 60
QY 634 TTLMLPCIGFAMLLMDLSGRFLLGTIPILIASLVILVSNLIDGLTAAHALSTVS 693
Db 61 TTLMLPSIGIAMRLMDSGRFLLLSTIPVLIALAVLVNVLVDGTMVHAALSTISV 120
QY 694 IYFCCFVMGFGPIPNILCAEIFPTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAIGLA 753
Db 121 IYFCFVMGFGPIPNILCAEIFPTSVRGICIAICALTFWIGDIIVTTLPMMLNAIGLA 180
QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLLEVITEFFAVGAKQ 795
Db 181 GVFGIYAVCVLAFVFMKVPETKGMPLLEVITEFFSVGAKQ 222

RESULT 15
US-10-051-909-16

; Sequence 16, Application US/10051909
; Publication NO. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-16

Query Match 23.1%; Score 943; DB 14; Length 228;
Best Local Similarity 80.6%; Pred. No. 6.6e-76;
Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 574 EPGVRRALLVGVGIQLIQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISL 633
Db 1 EPGVKHALFVIGLQILQFAGINGVLYTTPQILEQAGVAVILSNIGLSSSSASILISAL 60
QY 634 TTLMLPCIGFAMLLMDLSGRFLLGTIPILIASLVILVSNLIDGLTAAHALSTVS 693
Db 61 TTLMLPSIGIAMRLMDSGRFLLLSTIPVLIALAVLVNVLVDGTMVHAALSTISV 120
QY 694 IYFCCFVMGFGPIPNILCAEIFPTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAIGLA 753
Db 121 IYFCFVMGFGPIPNILCAEIFPTSVRGICIAICALTFWIGDIIVTTLPMMLNAIGLA 180
QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLLEVITEFFAVGAKQ 795
Db 181 GVFGIYAVCVLAFVFMKVPETKGMPLLEVITEFFSVGAKQ 222

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 19:05:50 ; Search time 4174 Seconds
(without alignments)
4658.261 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFPDLDLR.....PLEVITTEFAVGAKQAAMA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 segs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool/US10051909/runat_05012004_094742_22317/app_query.fasta_1.967
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909_CGN_1_1_2937 @runat_05012004_094742_22317 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vit: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pio: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vr1: *
28: gb_gss1: *

29: gb_gss2: *
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3027	74.3	2180	11 AY105508	AY105508 Zea mays
C 2	1399	34.3	911	29 CC423708	CC423708 PUH049TB
3	1382.5	33.9	845	29 BZ723844	BZ723844 PUCFI60TD
C 4	1279	31.4	795	29 BZ989776	BZ989776 PUGDY07TD
5	1101	27.0	829	13 BQ752737	BQ752737 WHE4118_E
6	990.5	24.3	745	14 CB635485	CB635485 OS1EB15O
C 7	981	24.1	594	9 AI861088	AI861088 603011H02
8	957	23.5	817	29 BZ819234	BZ819234 PUGDY07TB
9	930	22.8	859	12 BM817360	BM817360 HC106B05
10	928	22.8	749	13 BU040051	BU040051 PP_LBa000
11	924	22.7	688	14 CA502649	CA502649 WHE4338_A
12	921.5	22.6	700	13 BQ870435	BQ870435 QGD9A19_Y
13	921.5	22.6	703	14 CA502490	CA502490 WHE4048_B
14	916	22.5	847	14 CB981615	CB981615 CAB70005
15	909	22.3	572	12 BM325261	BM325261 PIC1_42_E
16	908	22.3	762	12 BI310999	BI310999 EST531274
17	902.5	22.1	710	13 BQ870328	BQ870328 QGD8L23_Y
18	895.5	22.0	738	14 CB965602	CB965602 NL23_F11
19	894	21.9	645	14 CA030046	CA030046 HX05020r
20	892.5	21.9	727	12 BJ471780	BJ471780 BJ471780
21	891	21.9	682	13 BU007863	BU007863 QGH5L06_Y
22	889	21.8	671	9 AV909064	AV909064 AV909064
23	888.5	21.8	776	14 CB671921	CB671921 OSJNE05L
24	877.5	21.5	739	12 BI920655	BI920655 EST540578
25	876	21.5	656	13 BU925675	BU925675 MCS015A08
26	875	21.5	688	13 BU004365	BU004365 QG4h09_Y
27	874.5	21.5	805	14 CB673272	CB673272 OSJNE07L
28	865	21.2	653	12 BJ450590	BJ450590 BJ450590
29	855	21.0	682	13 BU040035	BU040035 PP_LBa000
30	852	20.9	720	12 BI310710	BI310710 EST531246
31	850	20.9	819	10 BG584543	BG584543 EST486304
32	849.5	20.8	706	12 BJ472944	BJ472944 BJ472944
33	849	20.8	624	14 CB213154	CB213154 OML03434
34	847.5	20.8	787	14 CB634320	CB634320 OS1EB13M
35	840	20.6	617	13 BU991448	BU991448 HD06P09r
36	838	20.6	679	13 BU040591	BU040591 PP_LBa000
37	837.5	20.6	620	13 BU989641	BU989641 HF22E20r
38	836	20.5	609	12 BI309579	BI309579 EST530989
39	834.5	20.5	675	13 BU989051	BU989051 HF19J23r
40	833	20.4	725	13 BQ858206	BQ858206 QGB9K03_Y
41	831	20.4	627	14 CB816559	CB816559 3529_1_80
42	830.5	20.4	673	12 BJ449118	BJ449118 BJ449118
43	827	20.3	645	13 BQ240484	BQ240484 Tae05016G
44	825.5	20.3	751	13 BQ993492	BQ993492 QGF3f04_Y
45	824.5	20.2	699	14 CB966427	CB966427 NL37_F05

ALIGNMENTS

RESULT 1	AY105508	2180 bp	mrna	linear	HTC 16-OCT-2002
LOCUS	AY105508				
DEFINITION	Zea mays	PCO114533	mrna	sequence.	
ACCESSION	AY105508				
VERSION	AY105508.1	GI:21208586			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE	1	(bases 1 to 2180)			

AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2180)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Source Location/Qualifiers
1..2180
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:636667"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 533 a 447 c 555 g 645 t
ORIGIN
Alignment Scores:
Pred. No.: 5.93e-274 Length: 2180
Score: 3027.00 Matches: 591
Percent Similarity: 99.66% Conservative: 2
Best Local Similarity: 99.33% Mismatches: 0
Query Match: 74.28% Indels: 0
DB: 11 Gaps: 0
US-10-051-909-32 (1-800) x AY105508 (1-2180)
QY 206 MetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeu 225
DB 50 ATGGTGTGGGATGTCCTGTCGCCATCACCCGATGGAGAATATGCTGTGCTC 109
QY 226 AlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGlySerProArg 245
DB 110 GCGATACCTTCATGTTCTTCTTGGTTGACATATTTATCTTCTGAACTCTCCAGA 169
QY 246 TrpLeuValSerIleGlyArgMetAlaGluAlaLysValLeuGlnLysLeuArgGly 265
DB 170 TGGCTCGTTAGCAAGGTCGGATGGCAGAGGCAAAAGGTGTGCAAAAGTTACGGTGG 229
QY 266 LysAspAspValSerGlyGluLeuSerLeuLeuLeuGlnGlyLeuGlnValGlyLysAsp 285
DB 230 AAAGACGATGTCAGGTGAATGTCCCTTCTTCGAAAGGTTGAGGTTGAGAGAGAC 289
QY 286 ThrSerIleGluGluTyrIleIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThr 305
DB 290 ACTTCATGAGAGTACATCATGACCTGCCACGAGGACGCCGATGATCTTGTACT 349
QY 306 AspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAla 325
DB 350 GACGGTGATAAGAAACAATCACTTATGGGCTTGAAGAAGGCCAGTCATGATTGCT 409
QY 326 ArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGly 345
DB 410 CGACCTTCTAAGGAGCAACATCATGCTTGAAGTGTGCTTCTCTTGCATCTGTCATGGG 469
QY 346 SerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerVal 365

DB 470 AGCATGTGAACCAAGAGTGTACCCCTTATGATCCGATTGTGACACATTTTGTAGTGTGTC 529
QY 366 HisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGly 385
DB 530 CATGAGAATATGCTCAAGCTGAGAGAGATGAGAGACACATTTTCCAACTTTGGA 589
QY 386 SerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluAsnLeu 405
DB 590 AGTATGTTCAAGTGTCAAGATCAGATCAGATCCAAATAGACAGTGGATGAAGAATCTT 649
QY 406 HisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyLysAspTyrGluAspAsnLeu 425
DB 650 CATAGGATGACGAGAGTACGATCTGATGTCAGAGAGTGAATGACATCAATCTC 709
QY 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
DB 710 CATAGCCCATTTGCTGTCCAGGCGACGACACAGTGGGAAGGAGACATTTGTCCACCAT 769
QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyLysAsp 465
DB 770 GGTCAACCGTGAAGTGTCTTGAACATGAGAGGCAAAAGCTCTTAGGGGAGGCTGAGAT 829
QY 466 GlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGlyLys 485
DB 830 GGTGTGACGACACATGATATCGGTGGGAGTGCAGCTTGTGAAATGTCAGAGAAG 889
QY 486 GlnGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
DB 890 GAAGTGAGAAATGTGAAAGAAAGAGGTGTTCAAAGAGTCTACTTGCACCAAGAGGA 949
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyLysAspValPheGlu 525
DB 950 GTTCCTGGCTCAAGAGGGGCTCAATGTGTTCACTCCGGTGGTGGAGTCTCTTGAG 1009
QY 526 GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
DB 1010 GGTAGTGAGTTGTGATCATGCTGCTGCTTATGATGATCAGACACTTTCTCAAGGGT 1069
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
DB 1070 CTTCCTGAACCAACGATGATGATGCTGCCATGTTCAACCATCTGAGGTAGTCCCAA 1129
QY 566 GlySerArgTrpLysAspLeuPheGluProGlyValAlaArgAlaLeuLeuValGlyVal 585
DB 1130 GGTTCACGTTGAAAGATTGTTTGAACCTGAGTGAAGCGCTGCCCTGTAGTGGTGT 1189
QY 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTrpProGln 605
DB 1190 GGAATTCAAGATCCTTCAACAGTTGTGTAATAAAGGTGTCTGATCTAATACCCACAA 1249
QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
DB 1250 ATTCTGAGCAAGCTGTGTGGCAGTATTTCTTCCAATTGTGCTCAGCTCGGCATCA 1309
QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
DB 1310 GCATTCATCTTGATCAGTTCTCTCACTACCTTAATATGCTTCTTGCAATTGGCTTGGC 1369
QY 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeu 665
DB 1370 ATGCTGCTTATGATCTTCCGGAAGAGGTTTGTGCTAGGCACAATTCATCTTG 1429
QY 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
DB 1430 ATAGCATCTTAGTATCTCGTGTGTGCCAATGTAATTTGGTACACATGACCAT 1489
QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
DB 1490 GCTTGTCTCCACCATCAAGTGTATCGTCTACTCTGCTGCTTGTATGGAATTGGT 1549
QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725

Db 1550 CCCATCCCCACATTTTATGTGCAGAGATCTTCCACCAGGGTTCGTGCGCTCTGTATT 1609

QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745

Db 1610 GCCATTGTGCGCTTACATCTCGATCGAGATATCATGTCACCTACAGCCTTCTGTG 1669

QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765

Db 1670 ATGCTGAATGCTATGAGCTGGCGGGTGTTCAGCATATATGACATCGTATGCTTGATT 1729

QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785

Db 1730 TCCTTGTGTTCGTCTTCCTTAAGGTCCTCGAGACAAAGGGGATGCCCTTGAGGTTATT 1789

QY 786 ThrGluPhePheAlaValAlaGlyAlaLysGlnAlaAlaLysAla 800

Db 1790 ACCGAATCTTTGCAGTTGCTGCGAAGCAAGCGGCTGCAAAAGCC 1834

RESULT 2

CC423708/c 911 bp DNA linear GSS 19-MAY-2003

LOCUS PUHOR49TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBMTa499102,

DEFINITION genomic survey sequence.

ACCESSION CC423708

VERSION CC423708.1 GI:30903798

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 911)

REFERENCE

AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick ,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.

TITLE Maize Genomics Consortium

JOURNAL Unpublished

COMMENT Other_GSSs: PUHOR49TD

CONTACT: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

Source

1..911

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMBMTa499102"

/clone_lib="ZM_0.6_1.0_KB"

/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high COT selected genomic DNA library"

BASE COUNT 242 a 264 c 163 g 242 t

ORIGIN

Alignment Scores:

Pred. No.: 3.96e-121 length: 911

Score: 1399.00 Matches: 267

Percent Similarity: 98.90% Conservative: 2

Best Local Similarity: 98.16% Mismatches: 3

Query Match: 34.33% Indels: 0

DB: 29 Gaps: 0

US-10-051-909-32 (1-800) x CC423708 (1-911)

QY 270 SerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGlu 289

Db 818 ACAGGTGAATTGTCCCTTCTTCGAAAGGTTGAGGTTGAGAGACACTTCAATTGAA 759

QY 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLys 309

Db 758 GAGTACATCATTTGACCTGCCACCAGGACCCGATGATCTGTACTGACCGGTGATAAG 699

QY 310 GluGlnIleThrLeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLys 329

Db 698 GAACAATACACTTTATGGGCTGAAGAGCCAGTCATGATGATCTTGACCTTCCAAG 639

QY 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349

Db 638 GGACCCAGCATGCTTGAAGTGTGCTTCTTCTTGATCTCGTCATGGAGCATGTTGAAC 579

QY 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369

Db 578 CAGAGTATACCCCTTATGATCCGATGTGACACTTTTGTAGTGTCCATGAGAATATG 519

QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389

Db 518 CCTCAAGCTGGAGGAAGTATGAGAGCACATTGTTCCAACCTTGGAGATATGTTCAAGT 459

QY 390 ValThrAspGlnHisAlaLysAsnGluGlnThrAspGluGluAsnLeuHisArgAspAsp 409

Db 458 GTCACAGATCAGCATGCCAAAATGAGCAGTGGGATGAAGAGAATCTTCATAGCGATGAC 399

QY 410 GluGluTyrAlaSerAspGlyAlaGlyLysAspTyrGluAspAsnLeuHisSerProLeu 429

Db 398 GAGAGTACGCATCTGATGTGTCMAAGAGTGACTATGAGAGACAATCTCATAGCCCATTTG 339

QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisIleGlyHisArgGly 449

Db 338 CTGTCCAGGCAGGCAACAGGTGCGGAAGGAGAATGTTGTGACACCATGTGACCGTGGA 279

QY 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyLysAspGlyValSerSer 469

Db 278 AGTCTTTGAGCATGAGAAGGCAAGCCCTTAAAGGGAGGAGGTGAGATGCTGAGCAGC 219

QY 470 ThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGluAsn 489

Db 218 ACTGATATCGGTGGGGGATGGCAGCTTGCTTGGAATGTGTCAGAGAAGAGAGTGAAGAT 159

QY 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509

Db 158 GGTAGAAAGGAGGTGCTTCAAAAGAGCTTACCTGCACCAAGAGGAGTTCCTGGCTCA 99

QY 510 ArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPhe 529

Db 98 AGAAGGGGCTCAATTGTTCACTTCCCGGTGTGGCGATGTTCTTGAGGGTATGAGTTT 39

QY 530 ValHisAlaAlaLeuValSerGlnSerAlaLeu 541

Db 38 GTACATGCTGCTGCTTTAGTAAGTACAGTCAGCACTT 3

RESULT 3

BZ723844 845 bp DNA linear GSS 24-FEB-2003

LOCUS BZ723844

DEFINITION PUCFI60TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBMTa133J23,

genomic survey sequence.

ACCESSION BZ723844

VERSION BZ723844.1 GI:28519500

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 845)

REFERENCE

AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick ,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.

TITLE Maize Genomics Consortium

JOURNAL Unpublished

COMMENT Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Db 618 GAACAATCACACTTATGGCCCTGAAGAAGGCCAGTCATGATGCTCGACCTTCCAAG 559

QY 330 G1YProilemetleuGlySerValleuSerleuAlaSerArgHisGlySerMetValasn 349

Db 558 GGAACCCAGCATGCTTGGAAGTGTGCTTCTCTTGCACTCGTCATGGAGCATGTGTAAC 499

QY 350 GlnSerValProleuMetaspProileValThrleuPheGlySerValHisGluasnMet 369

Db 498 CAGAGTGTACCCCTTATGATGCCATTGTGACACCTTTTGTAGTGTCCATGAGAAATATG 439

QY 370 ProGlnAlaGlyGlySerMetArgSerThrleuPheProasnPheGlySerMetPheSer 389

Db 438 CCTCAAGCTGAGGAAGTATGAGAGACACATTGTTCCAACTTGGAAAGTATGTTCAAGT 379

QY 390 ValThrAspGlnHisAlaIysAsnGluGlnThrAspGluGluAsnLeuHisArgAspAsp 409

Db 378 GTCAAGATCAGCATGCGCAAAATGAGCAGTGGATGAGAAGAAATCTTCATAGGATGAC 319

QY 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProleu 429

Db 318 GAGGAGTACGCATCTGATGGTGACAGAGGTGACTATGAGCAATCTCCATAGCCCATTG 259

QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449

Db 258 CTGTCAGGCAGGCAACAGGTGCGGAAGGGAAGACATGTGCACCATGTCACCGTGA 199

QY 450 SerAlaIeuSerMetArgArgGlnThrleuLeuGlyGluGlyLysAspGlyValSerSer 469

Db 198 AGTGCTTTGAGCATGAGAGCAAAAGCCTCTTAGGGGAGGCTGAGATGATGACAGC 139

QY 470 ThrAspIleGlyGlyTyrGlnleuAlaTrrPlySerTrpSerGluLysGluGlyValasn 489

Db 138 ACTGATATCGGTGGGGATGCGACTTGCTTGAATGCTCAGAGAAGAGAGTGAGAAAT 79

QY 490 G1YArgLysGluGlyGlyPheLysArgValTyrleuHisGlnGluGlyValProGlySer 509

Db 78 GGTAAGAAAGGAGAGGTGTTCAAAAGACTCTACTTCACCAAGAGGAGTTCTTACTCA 19

QY 510 ArgArgGlySerIleVal 515

Db 18 AGAAGGGGCTCAATTGTT 1

RESULT 5

BQ752737 829 bp mRNA linear EST 23-JUL-2002

LOCUS WHE4118_E11_1222S wheat salt-stressed root cDNA library Triticum

DEFINITION aestivum cDNA clone WHE4118_E11_122, mRNA sequence.

ACCESSION BQ752737

VERSION BQ752737.1 GI:21930519

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 829)

AUTHORS Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K., Dvorak,J., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.

TITLE The structure and function of the expressed portion of the wheat genomes - Salt-stressed root cDNA library

JOURNAL Unpublished

COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: candersn@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: SK primer.

FEATURES

Location/Qualifiers

source

1. .829

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE4118_E11_122"

/tissue_type="Roots"

/dev_stage="Full tillering"

/lab_host="E. coli SOLR"

/clone_lib="wheat salt-stressed root cDNA library"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hydroponic plants grown to full tillering stage were treated with 150 mM NaCl for either 12 hours or 7 days. Root tissues of the plants subjected to both types of treatment were collected separately at University of California, Davis (E. Akhunov and K. Deal in J. Dvorak's lab). Total RNA was prepared separately from the two samples (12h and 7day treatments), and equal amount of RNA was then pooled. PolyA RNA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak) at the University of California, Davis. Colony plating, plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 177 a 199 c 215 g 238 t

ORIGIN

Alignment Scores:

Pred. No.: 4.09e-93 Length: 829

Score: 1101.00 Matches: 220

Percent Similarity: 89.89% Conservative: 20

Best Local Similarity: 82.40% Mismatches: 25

Query Match: 27.02% Indels: 2

DB: 13 Gaps: 2

US-10-051-909-32 (1-800) x BQ752737 (1-829)

QY 486 GluGlyLysAsnGlyArgLysGluGlyPheLysArgValTyrleuHisGlnGluGly 505

Db 28 AAAGCGAGGATGCGAAGAAGAGAGGCTTCAAAAGATCTACTTGCACCAAGAGGGG 87

QY 506 ValProGlySerArgArgGlySerIleValSerleuProGlyGlyLysAspValPheGlu 525

Db 88 GTGGCCGACTCAAGAGGGGCTGTGTTCACTTCTGTTGGGGGTATGCCACCCAA 147

QY 526 --GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLys 544

Db 148 GGGGCGAGTGGGTTTATACACGCTGCTGTGTAAGCCACTCGGCTCTTACTCCAAG 207

QY 545 G1YLeuAlaGluProArgMet--SerAspAlaIleMetValHisProSerGluValAla 563

Db 208 GATCTTATGGAAGCGTATGCGCGCGTCAGCCATCATTCATTCGAGGACAGCT 267

QY 564 AlAlysGlySerArgTrpLysAspPheGluProGlyValArgArgAlaLeuLeuVal 583

Db 268 CCCAAGGTTCAATCTGGAAGAAGATCTGTTGAACCTGTTGAGGCGTGCAATTGTCGTC 327

QY 584 G1YValGlyIleGlnleuLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTrpThr 603

Db 328 GGGCTTGGAATTCAGATGCTTCAAGCAAGTTGCTGGAATAATGAGATTCTTACTATACT 387

QY 604 ProGlnleuLeuGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623

Db 388 CCTCAAAATTCTGAGCAAGCTGTGTGGCGGTTCTTCTTCCAATCTTGCCCTCAGTTCA 447

QY 624 AlaSerAlaSerIleLeuIleSerSerleuThrThrleuLeuMetleuProCysIleGly 643

Db 448 GCATCAGCGTCAATCTGATCAGCTCTCTCACCACTTACTCATGCTCCAGCAATTGGT 507

QY 644 PheAlaMetleuLeuMetAspLeuSerGlyArgArgPheleuLeuGlyThrIlePro 663

Db 508 GTAGCCATGAGACTTATGATATATATCTGAAGAAGGTTTCTTCTACTGGGCACAATTCCC 567

QY 664 IleuIleAlaSerLeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeu 683
Db 568 ATCTGATAGCATCCCTAATGTTGGTGTGTCCATGTTATCACTTGAGTACGGTG 627
QY 684 AlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGly 703
Db 628 CCCGACGCTGTCTCTCCACAGTTAGCGTCATTGCTTACTTCTGCTGCTTGTGATGGGC 687
QY 704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723
Db 688 TTGGCCCAATCCCAACATTCATGTGAGATTTCGCCACAGAGTCCGTTGGTGC 747
QY 724 CysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeu 743
Db 748 TGCATGCTATTGCGCCCTCACACTTGATATGTACATTAATTGTTACCTACAGCCTG 807
QY 744 ProValMetLeuAsnAlaIle 750
Db 808 CCTGTATGCTGATGCTATT 828

RESULT 6
CB635485 745 bp mRNA linear EST 08-APR-2003
LOCUS OSIEB15024.f OSIEB Oryza sativa (indica cultivar-group) cDNA
DEFINITION clone OSIEB15024 5', mRNA sequence.
ACCESSION CB635485
VERSION CB635485
KEYWORDS CB635485.1 GI:29630476
SOURCE EST.

ORGANISM Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erbartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 745)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: 0 column: 24
Seq primer: gta aac cga cgg cca gtcg.

FEATURES
Source
1..745
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIEB15024"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIEB"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"
BASE COUNT 165 a 161 c 216 g 203 t
ORIGIN

Alignment Scores:
Pred. No.: 9.02e-83 Length: 745
Score: 990.50 Matches: 192
Percent Similarity: 85.94% Conservative: 22

Best Local Similarity: 77.11% Mismatches: 34
Query Match: 24.31% Indels: 1
DB: 14 Gaps: 1
US-10-051-909-32 (1-800) x CB635485 (1-745)

QY 131 LeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsn 150
Db 2 CTGATCGCGTCCCTGCTCTACTTGTAGTGGCTAGTATGCTTTGGGCCCAAT 61
QY 151 ValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170
Db 62 GTGTATGTGTGCTTGGCGAGGCTCATTTGACGGGTTCCGGATCCGTTGGCTGCACG 121
QY 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190
Db 122 CTGTACCATTTGACATCTCTGAGACTGCCCGACGACATCAGAGACTGTAAACACG 181
QY 191 LeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet 210
Db 182 CTGCCGAGTTACGTGGTCTGAGAGGAGTTCCTTTCATACCTGACATGTAATTGGCATG 241
QY 211 SerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeu 230
Db 242 TCCCTCATGCCACAGCCAGATTGGAGATCATGCTTGGCGTTCTATCATCATCACTT 301
QY 231 PhePhePheGlyLeuThrIlePheTyrIleProGlnSerProArgTrpLeuValSerLys 250
Db 302 ATATACTTTCATGACCATCTTTACTTACTTACCTGAATCGCCGAGTGGCTGAGCAAA 361
QY 251 GlyArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspValSer 270
Db 362 GGAAGATGGCTGAGGCCCAAGCGTGTGTGCAAGCCCTGCGTGAAGAGACATGTTTCA 421
QY 271 GlyGluLeuSerLeuLeuLeuGlyGlyLeuGluValGlyLysAspThrSerIleGlu 290
Db 422 GAGAAATAGCCCTTCTCTGTGAAGGTCTGGGGCTGGAAAGACACAAATTGAGGAA 481
QY 291 TyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlu 310
Db 482 TACATAATTGACCTGATGATGAGCTTGTGATGAAGCGCTGCTCA---GATCCAGAG 538
QY 311 GlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGly 330
Db 539 AAGATCAACTGTATGCTCTGAAGAAGGCTTATCGTGGTGGCCGCTGTTCAACGGG 598
QY 331 ProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGln 350
Db 599 CAAAGTGACCTTGAAGTGCAATTAGGTCTCATCTCTGTCATGTATGTGTCAGTCA 658
QY 351 SerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPro 370
Db 659 GGTAGCCCCCTTGTGATCTGTGTACCCCTTTTGAAGTGTCCATGAGAACATGCT 718
QY 371 GlnAlaGlyGlySerMetArgSerThr 379
Db 719 GAGATAATGGAGCATGCGGAGCACA 745

RESULT 7
AI861088 594 bp mRNA linear EST 19-JUL-1999
LOCUS 603011H02.x1 603 - stressed root cDNA library from Wang/Bohnert lab
DEFINITION Zea mays cDNA, mRNA sequence.
ACCESSION AI861088
VERSION AI861088.1 GI:5525249
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 594).
AUTHORS Walbot,V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603011 row: H column: 02.
Location/Qualifiers

FEATURES
source
1. 594
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/note="Organ: root; Vector: pBluescriptII SK(+); XR; Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 210 a 127 c 142 g 115 t

ORIGIN

Alignment Scores:
Pred. No.: 5.13e-82 Length: 594
Score: 981.00 Matches: 197
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.07% Indels: 0
DB: 9 Gaps: 0

US-10-051-909-32 (1-800) x A1861088 (1-594)

QY 596 IleaNGlyValLeuTYrTYrThrProGlnIleLeuGlnAlaGlyValAlaValIle 615
Db 592 ATRAACGGTGTCTGTACTATACCCACAATCTTGAGCAAGCTGTGTGAGTTATT 533
QY 616 LeuSerLySPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 635
Db 532 CTTTCCAAATTGGTCTCAGCTGGCATCAGCATCCATCTGATCAGTTCTCTCATTACC 473
QY 636 LeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArg 655
Db 472 TTAATAATGCTTCTGCTGCTTGGCCATGCTGCTATGATCTTCCGGAAGAAG 413
QY 656 PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValIleSer 675
Db 412 TTTTGTGCTAGGACACAATTCATCTTGATAGCATCTTACTTATCTGTTGTCTCC 353
QY 676 AsnLeuIleAspLeuGlyThrIleAlaHisAlaLeuLeuSerThrValSerValIleVal 695
Db 352 AATCTAATTGATTGGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATCTGC 293
QY 696 TyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIle 715
Db 292 TACTTGTGCTGCTTGTATGGGATTGTGCTCCATCCCAACATTTTATGTGACAGATC 233
QY 716 PheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGly 735
Db 232 TTTCCACACAGGGTTCGTGGCTCTGTATTGCCATTTGTGCTTACATTCTGATCGGA 173
QY 736 AspIleIleValThrTyrSerIleProValMetLeuAsnAlaIleGlyLeuAlaGlyVal 755
Db 172 GATATCATCGTCACCTACAGCCTTCTGTGATGCTGATGCTATGTGACTGGCGGTGT 113
QY 756 PheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLyValPro 775
Db 112 TTCAGCATATATGACAGTCGATGCTTGTATTTCTTGTGTCTTCTTAAAGGTCCCT 53

QY 776 GluThrLySGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
Db 52 GAGACAAAGGGATGCCCCCTTGAGGTATTAACGAATCTTGTGAGTTGCT 2

RESULT 8
BZ819234 817 bp DNA linear GSS 18-MAR-2003
LOCUS
DEFINITION PUGDY07TB ZM_0.6_1.0_KB Zea mays genomic clone ZMBMTa362B13,
genomic survey sequence.
ACCESSION BZ819234
VERSION BZ819234.1 GI:29034056
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 817)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. 817
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

BASE COUNT 232 a 145 c 203 g 237 t

ORIGIN

Alignment Scores:
Pred. No.: 1.47e-79 Length: 817
Score: 957.00 Matches: 183
Percent Similarity: 99.46% Conservative: 1
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 23.48% Indels: 0
DB: 29 Gaps: 0

US-10-051-909-32 (1-800) x BZ819234 (1-817)

QY 270 SerGlyGluLeuSerIleLeuLeuGluGlyLeuGluValGlyIleAspThrSerIleGlu 289
Db 262 ACAGGTGAATTGTCTTCTTCTCGAAGGGTTGAGGTTGAGGAGACACTTCCATTGAA 321
QY 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThraSpGlyAspLys 309
Db 322 GAGTACATCATTTGACCTGACCCAGGACGAGCATGATCTTGTACTGACGGTGAAGA 381
QY 310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329
Db 382 GAACAATCACACTTATGGGCTGAAGAAGGCCAGTCATGATGTGCTGACTTCCAAG 441
QY 330 GlyProIleMetLeuGlySerValLeuSerIleAlaSerArgHisGlySerMetValAsn 349
Db 442 GACACCAAGCATGCTTGAAGTGTGCTTCTTCTGATCTCGTCATGGAGCATGTGAAC 501
QY 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369
Db 502 CAGAGTGACCCCTTATGATCCGATTGTGACATTTTTGTAGTGTCCATGAGAATATG 561

QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPhGlySerMetPheSer 389
Db 562 CCTCAAGCTGAGGAGAGTATGAGGACATGTTTCCAACTTTGGAAGTATGTTCACT 621
QY 390 ValThrAspGlnHisAlaLysAsnGlnTrpAspGluGluAsnLeuHisArgAspAsp 409
Db 622 GTCACAGATCAGCATGCCCCAAAATGACAGTGGATGAAGAGATCTTCATAGGATGAC 681
QY 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429
Db 682 GAGAGTACGCATCTGATGTGACAGAGTGACTATGAGACATCTCCATAGCCCCATG 741
QY 430 LeuSerArgGlnAlaThrGlyAlaGlyGlyLysAspIleValHisHisGlyHisArgGly 449
Db 742 CTGTCCAGGACGACCAACAGGTCCGGAAGGACCATTTGTGCACCATGTCACCCGTGA 801
QY 450 SerAlaLeuSerMet 454
Db 802 AGTCTTTGAGCATG 816
RESULT 9
LOCUS BM817360 859 bp mRNA linear EST 05-MAR-2002
DEFINITION HC106B05_T3.ab1 HC Hordeum vulgare subsp. vulgare cDNA clone
HC106B05_T3.ab1 similar to (AC007369) Sugar transporter
[Arabidopsis thaliana], mRNA sequence.
ACCESSION BM817360 GI:19153374
VERSION BM817360.1 GI:19153374
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 859)
REFERENCE
AUTHORS Ozturk,N.Z., Michalowski,C.B., Brazille,S., Borchert,C., Palacio,C.,
Noemand,C., Murphy,C., Kelley,R., Sant,S.A., McLaughlin,H.,
Fredricksen,M.A. and Bohnert,H.J.
TITLE Monitoring large-scale changes in transcript abundance in drought-
and salt-stressed barley
JOURNAL Unpublished (2002)
COMMENT Contact: Mark A. Fredricksen
Plant Biology
University of Illinois
1201 W Gregory Dr, Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
FEATURES
source
1. 859
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/strain="cv tokak"
/db_xref="taxon:112509"
/clone="HC106B05_T3.ab1"
/issue_type="Root"
/dev_stage="3 week old"
/clone_lib="HC"
/note="6 and 10 hour drought stress by placing plants on
moist paper (75% rel. humidity) in light"
BASE COUNT 179 a 206 c 213 g 260 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 5.53e-77 length: 859
Score: 930.00 Matches: 183
Percent Similarity: 80.58% Conservative: 41
Best Local Similarity: 65.83% Mismatches: 51
Query Match: 22.82% Indels: 3
DB: 12 Gaps: 1
US-10-051-909-32 (1-800) x BM817360 (1-859)
QY 496 PheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleVal 515

Db 34 TTCCCAAGGATTTATTTCATCAGGAAGGTGAGCTTGATCAAAACGAGGCTTATATT 93
QY 516 SerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeu 535
Db 94 TCTCTT-----GGTGTGATGCTGCTGAGAGAACGAGTATATCAAGGCTGTGACTA 147
QY 536 ValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 555
Db 148 GTAAGTCAACGAGCCCTTATTTCAGAGTGTATGATCGGAATTCCGTTGACAGCT 207
QY 556 MetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluPro 575
Db 208 ATGATTCAACCCCTGCAAGCTCTGCTAAAGGCTTATCTGAGGATCTGTTGAGCCT 267
QY 576 GlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGly 595
Db 268 GGTGTCAAGCATGCATTTGGTTGGTGTGGTGTGGACTTCAATTCTTCAGCAGTTCTCAGGG 327
QY 596 IleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlu-GlnAlaGlyValAlaValIle 615
Db 328 ATAGCGGAGTTATGATTAATACTCCCAATACTTCAGCCAGGCTGGGGTGTGATCT 387
QY 615 eleuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeuThrTh 635
Db 388 TCTATCACACATAGGATTAAGCCAGTCATCTGCGTCAATTCCTCATTAAGTCACTGACGAC 447
QY 635 rleuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgAr 655
Db 448 ACTTCTGATGCTTCTCTATAGCTGTGCTATGAGGCTCATGATATATCAGAAAGAG 507
QY 655 gPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValIse 675
Db 508 GAGTTTGCTGCTAAGACATCCCAATGTGATATCTCCCTCGTTGCTCATCGG 567
QY 675 rAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValIleLeuVal 695
Db 568 AAGCTTAATAAAGATGGATCCACTATAATAGCCGTGATCTCCACTGTATGTGTGCT 627
QY 695 lTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluI 715
Db 628 TTACTTCTGTTCTTCTGATGAGGCTTTGGCCGCTTCCCAACATCTCTGCTCGAGAT 687
QY 715 ePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleG 735
Db 688 CTTCCNCACCCGTTGCTGCGGCTGATGCCATTGTGACATCACTTTTGATTAG 747
QY 735 yAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyVa 755
Db 748 TGACATCATGCTGACCTACACACTCCCGTATGCTCAATCTTTGGAATCGCTGGCT 807
QY 755 lPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db 808 CTTTGGTATCTATGGCGTGGGTGCAATAACTCATGGGTTTTCATCTTCTT 859
RESULT 10
LOCUS BU040051 749 bp mRNA linear EST 26-AUG-2002
DEFINITION PP_LBa0004L05f Peach developing fruit mesocarp Prunus persica cDNA
clone PP_LBa0004L05f, mRNA sequence.
ACCESSION BU040051 GI:22478805
VERSION BU040051.1 GI:22478805
KEYWORDS EST.
SOURCE Prunus persica (peach)
ORGANISM Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
1 (bases 1 to 749)
REFERENCE
AUTHORS Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
TITLE Peach Model Genome for Rosaceae
JOURNAL Unpublished
COMMENT Contact: Abbott, A.

Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 604
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 749.
Location/Qualifiers

FEATURES

1. 749

/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEA0004L05f"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site_1: EcoRI;
Site_2: XhoI; authority=Prunus persica L. Batsch; The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
http://www.genome.clemson.edu/projects/peach. To order
this clone go to http://www.genome.clemson.edu/orders"
BASE COUNT 167 a 167 c 163 g 252 t
ORIGIN

Alignment Scores:

Pred. No.: 7.02e-77 length: 749
Score: 928.00 Matches: 172
Percent Similarity: 87.23% Conservative: 33
Best Local Similarity: 73.19% Mismatches: 30
Query Match: 22.77% Indels: 0
DB: 13 Gaps: 0

US-10-051-909-32 (1-800) x BU040051 (1-749)

QY 564 AlAlYSGLYSerArgTTrpYsAsPLeuPheGluProGlyValArgAlaLeuLeuVal 583
Db 3 TCAGAAGGACCAATCTGGGCTGCTCTGTTGAACCAAGGGTTAAGCATGCTTTCGTT 62
QY 584 GLYValGLYlleGlnlleGlnGlnPheAlaGlylleAsnGlyValLeuTyrThr 603
Db 63 GGAATAGGAATCCAGATTCTTCAGCAGATTTCGTGGCAATATGGGGTCTGTACTACACT 122
QY 604 ProGlnlleGlnGlnGlnGlnPheAlaGlylleAsnGlyValLeuTyrThr 623
Db 123 CCTCAAAATTCCTGAAGATGCAAGTGTGAAGTCTTCTTGAAGAAGCTTGGGTCTCAGTACA 182
QY 624 AlAserAlaserlleleuIleserSerleuThrThrleuLeuMetleuProCyslleGly 643
Db 183 GAGTCTTCATCCTCTCAATAGTGCAATTCACAACTTGTGTGATGCTTCTGTATAGCC 242
QY 644 PheAlaMetleuLeuMetAsPLeuSerGlyArgArgPheLeuLeuGlyThrIlePro 663
Db 243 ATAGCCATAAGCTCATGATATCTCTGTAGAGAGCGCTGCTACTGGCTACAAATTCCT 302
QY 664 lleleuIleAlaserlleleuValIleValSerAsnleuIleAsPLeuGlyThrIleu 683
Db 303 GTGTGTGATGACACTCATCATCTTAATCATGCGCACTTGTGACCTTGGGTACAGTT 362
QY 684 AlaHisAlaLeuLeuSerThrValSerValIleValIyrPheCysCysPheValMetGly 703
Db 363 CTTCAATGCGGCTTATCATGCTATTGTGTATATGTCATATTCTGCTGCTTGTATGCGCC 422
QY 704 PheGlyProIleProAsnIleleuCysAlaGluIlePheProThrArgValArgGlyLeu 723
Db 423 TATGGGCCAAATTCCAATATCTCTGCTCAGAGATCTTCCGACAAAGGTACGTGCGCTC 482
QY 724 CysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerleu 743

Db 483 TGCATTGCCATCTGCGCCCTAGTGTACTGGAATTGGGACATTAATGTCACCTACACTA 542
QY 744 ProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValCys 763
Db 543 CCTGTGCTGCTTGAATTCATATAGCCCTAGCTGTATCTTGTGATTTATGTGTGTTGT 602
QY 764 leuIleSerPheValPheValPheleuIleValProGluThrIleGlyMetProleuGlu 783
Db 603 GTCATTTCCTTTATTTATCTTCTTGAAGTTCCAGAAACCAAGGCAATGCCCTTGAA 662
QY 784 ValIleThrGluPhePheAlaValGlyAlaIleAlaIleAlaIleAlaIleAlaIle 798
Db 663 GTCATTACTGAATCTTGTGCTGCGTGACAGACAAGTTGCCGCA 707

RESULT 11

CA502649 688 bp mRNA linear EST 14-NOV-2002
LOCUS WHE4338_A09_A18ZS wheat meiotic anther cDNA library Triticum
DEFINITION aestivum cDNA clone WHE4338_A09_A18, mRNA sequence.
ACCESSION CA502649
VERSION CA502649.1 GI:24993603
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticaceae; Triticum.
1 (bases 1 to 688)
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Iazo, G.R., Pham
J., Rausch, C.J., Sutton, T., Woo, J., and Wilson, C.
The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
Unpublished
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES

Location/Qualifiers
1. 688
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4338_A09_A18"
/tissue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="Wheat meiotic anther cDNA library"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants
were grown in a glasshouse. Anther meiotic stage was
determined by removing anthers from individual primary
flowers. One anther was sacrificed for microscopic staging
, and if determined to be between (and including) meiotic
stages pre-meiosis and metaphase I, the remaining two
anthers were collected and pooled for library
construction. The tissue, total RNA, and poly(A) RNA were
prepared, cDNA synthesised, and directionally ligated into
pSPORT1 by Tim Sutton in the P Langridge lab at the
Department of Plant Science, University of Adelaide, Waite
Campus, Australia. Average insert size 1.5Kb. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."

BASE COUNT 138 a 167 c 165 g 218 t
ORIGIN

Alignment Scores:

Pred. No.: 1.48e-76 Length: 688
Score: 924.00 Matches: 179
Percent Similarity: 87.34% Conservative: 21
Best Local Similarity: 78.17% Mismatches: 29
Query Match: 22.67% Indels: 0
DB: 14 Gaps: 0

US-10-051-909-32 (1-800) x CA502649 (1-688)

QY 538 GlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetVal 557
DB 2 CAACCAGCTCTTACTCAAGGAAGTGTGAGCAACGCTGCTGCTCTGCGATGATG 61
QY 558 HisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyVal 577
DB 62 CATCCATCTGAGGAGTGTGCTAAAGGTCCAAGATG3GCTGACCTGTTGAGCCTGAGATG 121
QY 578 ArgArgAlaLeuLeuValGlyValGlyLeuGlnLeuGlnPheAlaGlyIleAsn 597
DB 122 AAGCATGCACTGTTGTTGGCATAGATACAGATCCTGCACAGCTTGTGCGCATCAAT 181
QY 598 GlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSer 617
DB 182 GGTGTCTCTACTACACTCCTCAAAATCTTGAGCAAGCAGGTGTGCTTCTTCTGTGC 241
QY 618 LysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeu 637
DB 242 AAGATTGGCTTAGCGCATCTTCTGCAATCTTATTAGTCTTGACAACTTATTG 301
QY 638 MetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeu 657
DB 302 ATGCTTCAAGCATTTGATTTGCAATGAGGCTCATGATATGCTGGAAGAGATTCTT 361
QY 658 LeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeu 677
DB 362 CTGCTTGGACAAATCCCTATCTGATAGTTGCCCTAGCTATCTGTGCTGCAATATT 421
QY 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
DB 422 GTGATGTGGGAACCATGTGTCATGCTGCACTCTCCACGATTAGTGTATTTTC 481
QY 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717
DB 482 TGGTCTTTTGTATGGGGTTGGCTATTCCCAACATTCTGTGTGCAAGATCTTTCC 541
QY 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
DB 542 ACCACCTCCGGCGCATCTGCATAGCCATCTGCCCTTAACCTTCTGATTGTTGACATT 601
QY 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
DB 602 ATGCTGACATACACACTTCTGTGATGCTGAATGCATCGGCTCGCTGTCTTTGGG 661
QY 758 IleTyrAlaValValCysLeuIleSer 766
DB 662 ATATTACGCCGCTGTTGCATCTGCGCT 688
RESULT 12
BQ870435 700 bp mRNA linear EST 15-AUG-2002
LOCUS BQ870435
DEFINITION OGD9A19.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone
OGD9A19, mRNA sequence.
ACCESSION BQ870435
VERSION BQ870435.1 GI:22256844
KEYWORDS EST.
SOURCE Lactuca sativa
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.
REFERENCE 1 (bases 1 to 700)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,

TITLE
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aamundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
singleton, see http://cgpdb.ucdavis.edu/ for details.
Plate: QGD9 row: A column: 19.
Location/Qualifiers

FEATURES
source

1. 700
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGD9A19"
/lab_host="E.coli"
/clone_lib="QG_ABCDI lettuce salinas"
/note="Vector: pBRCDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and size bias details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/TAG_SEQ=Not found"

BASE COUNT 166 a 136 c 175 g 223 t
ORIGIN

Alignment Scores:
Pred. No.: 2.61e-76 Length: 700
Score: 921.50 Matches: 174
Percent Similarity: 89.13% Conservative: 31
Best Local Similarity: 75.65% Mismatches: 22
Query Match: 22.61% Indels: 3
DB: 13 Gaps: 1

US-10-051-909-32 (1-800) x BQ870435 (1-700)

QY 80 IleAlaAlaValLeuTyrIleLysGlyPheGlnLeuGlnAsnGluProThrVal 99
DB 20 GTAGCCGGGCTTTTATACATAAAGAAGATTCAATTGCAAAACGAACCAACGATC 79
QY 100 GluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGly 119
DB 80 GAGGCTTAATTGTGCTATGTGCTGATCGGTGCGACTTATGACGACTTGCTCCGGT 139
QY 120 ProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhe 139
DB 140 GCAATATCAGATTGGGTGGCGCGACGTCGATCATCTCTTCTGTTATTATT 199
QY 140 PheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPhe 159
DB 200 TTTGGTGTCTTGTATGTTATGATGATCAACCAATGTTATATCTTTATTTGAGTACTA 259
QY 160 ValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIle 179
DB 260 TTAGATGTTTGAATCGGATTAGCGGTACTCTGTTCCCGTTTACATATCGGAACA 319
QY 180 AlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGly 199
DB 320 TCACCGCGGAAATCAGAGGTTCACTCAACACTTTACACAGATTGAGGCTGTGGA 379
QY 200 MetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArg 219
|||||

[illegible]

BASE COUNT	ORIGIN
139 a	Campus, Australia. Average insert size 1.5Kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors). "
183 c	
165 g	
216 t	

Alignment Scores:	
Pred. No.:	2.63e-76
Score:	921.50
Percent Similarity:	86.70%
Best Local Similarity:	81.12%
Query Match:	22.61%
DB:	14
	Gaps: 2
	Length: 703
	Matches: 189
	Conservative: 13
	Mismatches: 24
	Indels: 7
	Gaps: 2

US-10-051-909-32 (1-800) X CA502490 (1-703)

Qy	526	GLYSerGluPhe-----ValHisAlaAlaLeuValSerGlnSer	539
Db	3	GGTCCGGAATTTCCGGGTGCACCACGCGTCCGACACGACTGCTTTGGTAAAGCCACTCG	62
Qy	540	AlaLeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHis	558
		:::	:::
Db	63	GCTCTTACTCCAGGATCTTATGGAAGAGCGGTATGGCGGCTGTGCCAGCATGATTTCAT	122
Qy	559	ProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArg	578
Db	123	CCATTGGAGGCGAGCTCCCAAGGTTCAATCTGGAAGAATCTGTTGAACCTGGTGTAGG	182
Qy	579	ArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleasnGly	598
Db	183	CGTGCATTGTTGTCGGCGGTTGGAATTCAAGATGCTTCACAGATTGCTGGAATAAATGA	242
Qy	599	ValLeuTyrrTrpThrProGlnIleLeuGlnGlnAlaGlyValAlaValIleLeuSerLys	618
Db	243	GTTCTCTACTATACTCTCAAAATTCTGGAGCAAGCTGTGTGGCGTTCTTCTTCCAAAT	302
Qy	619	PheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMet	638
Db	303	CTTGGCCTCAGTTCAGATCAGCATCCATCTTGATCAGTTCTCTCACCACCTTACTCATG	362
Qy	639	LeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu	658
Db	363	CTCCCAAGCATTGGTGTAGCCATGAGACTTAIGATATATCTGGAAGAAAGGTTTCTGCTA	422
Qy	659	LeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIle	678
Db	423	CTAGGCACAATTCCTCAATCTTGATAGCATCCCTTAATGTTTGGTGTGCCAATGTTATC	482
Qy	679	AspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCys	698
		::: :::	:::
Db	483	AAC TTGAGTACGGGTGCCCCACGCTGTGCTCTCCACAGTTAGCGTCATTGCTACTTCTGC	542
Qy	699	CysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThr	718
Db	543	TGCTTTGTCATGGGCTTTGGCCCCGATCCCCACAATCTATATGTGCAGAGATTTCGCCACC	602
Qy	719	ArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIle	738
		:::	:::
Db	603	AGAGCCCGGTGTGTGTCATCGCTATTGTGGCCCTCACATTCGTGATTGTGACATTATT	662
Qy	739	ValThrTyrrSerLeuProValMetLeuAsnAlaIleGly	751
Db	663	GGTACCTACAGCCTGCCTGTGATGCTGAATGCTATTGTT	701
RESULT 14			
CB981615			
LOCUS	CB981615	847 bp	mRNA linear EST 01-MAY-2003
DEFINITION	CAB70005.1 iaf_A05 Cabernet Sauvignon Berry Post-Veraison - CAB7		
	vitis vinifera cDNA clone CAB70005_1 iaf_A05 5', mRNA sequence.		
ACCESSION	CB981615		
VERSION	CB981615.1 GI:30304821		
KEYWORDS	EST.		
SOURCE	Vitis vinifera		

/note="Vector: pBluescript II SK(-) from Lambda Zap II;
Site 1: XhoI; Site 2: EcoRI; Four-week-old sorghum seedlings were sprayed with spore suspension prepared from 3-week-old FRM42I, a sorghum isolate of the antheranose pathogen Colletotrichum graminicola. Inoculated plants were kept in a 25 C dark growth chamber with 100% relative humidity for 24 hr, followed by 12/12 hr of light/dark cycle at 25 C with 90% relative humidity for another 24 hr. All leaves were harvested and quick frozen with liquid nitrogen and stored in a -80 C freezer. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision.
WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 134 a 124 c 152 g 162 t
ORIGIN

Alignment Scores:
Pred. No.: 2.95e-75 Length: 572
Score: 909.00 Matches: 177
Percent Similarity: 96.26% Conservative: 3
Best Local Similarity: 94.65% Mismatches: 7
Query Match: 22.31% Indels: 0
DB: 12 Gaps: 0

US-10-051-909-32 (1-800) x BM325261 (1-572)

QY	141	SerglyLeuIleMetLeuTrpSerProAsnValTyValLeuLeuLeuAlaArgPheVal	160
DB	12	AGCAGCCTCATCATGCTATGGTCCCTAATGTCATATGCTCTGCTGCGACGCTCGTA	71
QY	161	AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyriLeSerglyLeuAla	180
DB	72	GACGGATTGGVATGGCTTGCTGTCACGCTTGCTTGACATTTCAGAAACAGCC	131
QY	181	ProSerglyLeuIleArgGlyLeuLeuAsnThrLeuProGlnPheSerglySerglyMet	200
DB	132	CCTCCAGAGATTAGAGGTTGGTGAATACACTGCCACAGTTCAAGTATCAGAGGATG	191
QY	201	PheLeuSerTyrcysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle	220
DB	192	TTCTTGTCATACATGATGGTGTGGGATGTCACTGTGCCATCACCTGATTGAGAAATT	251
QY	221	MetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrlleu	240
DB	252	ATGCTTGGGGTGCTCGCGATACCTTCAATTGTTCTTGGATTGACAATATTTACCTT	311
QY	241	ProGluSerProArgTrpLeuValSerlySGlyArgMetAlaGluAlaLysLysValleu	260
DB	312	CCTGAATCCCAGAGATGGCTTGTAAGCAAGGTCGATGCGCAGGCAAGAGGTGTG	371
QY	261	GlnLysLeuArgGlyLysAspAspValSerglyGluLeuSerLeuLeuGluGlyLeu	280
DB	372	CAAAAATTACGACGCAAGAAAGATGTCTCAGGTGAATTGTCCTTCTTGAAGGTTG	431
QY	281	GluValGlyGlyAspThrSerIleGluGluTyriIleIleGlyProAlaThrGluAlaAla	300
DB	432	GAGGTGGAGAGACACTTCGATTGAAGATACATCATTTGCCCTGCCACTGACGCGCC	491
QY	301	AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrglyProGluGluGly	320
DB	492	GATGATCATGTTACTGATGGTGAAGAAACAATCACACTTTATGGGCTGAAGAAAGGC	551
QY	321	GlnSerTrpIleAlaArgPro	327
DB	552	CAGTCATGGATTGCTGACCT	572

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 21:45:08 ; Search time 648 Seconds
(without alignments)
4273.180 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPLDRR.....PLEVITFEFAVGAKQAATA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2263443 segs, 1730637950 residues
Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool/US10051909/runat_05012004_094745_22401/app_query.fasta_1.967
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10051909 @CGN 1 1 418 @runat_05012004_094745_22401
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2: *
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq: *
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq: *
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	4075	100.0	2777	14	US-10-051-909-31	Sequence 31, Appl
2	4059	99.6	2908	12	US-10-310-154-350	Sequence 350, App
3	3517	86.3	2824	14	US-10-051-902-1	Sequence 1, Appli
4	3517	86.3	2824	14	US-10-051-909-1	Sequence 1, Appli
5	2674	65.6	2601	14	US-10-051-902-7	Sequence 7, Appli
6	2674	65.6	2601	14	US-10-051-909-7	Sequence 7, Appli
7	2584	63.4	2190	10	US-09-938-842A-1315	Sequence 1315, Ap
8	2275	55.8	2205	10	US-09-938-842A-2254	Sequence 2254, Ap
9	1670.5	41.0	1692	14	US-10-051-902-9	Sequence 9, Appli
10	1670.5	41.0	1692	14	US-10-051-909-9	Sequence 9, Appli
11	1424	34.9	1487	14	US-10-051-902-13	Sequence 13, Appl
12	1424	34.9	1487	14	US-10-051-909-13	Sequence 13, Appl
13	943	23.1	1009	14	US-10-051-902-15	Sequence 15, Appl
14	943	23.1	1009	14	US-10-051-909-15	Sequence 15, Appl
15	587.5	14.4	1374	12	US-10-369-493-47011	Sequence 47011, A
16	562	13.8	870	14	US-10-051-902-5	Sequence 5, Appli
17	562	13.8	870	14	US-10-051-909-5	Sequence 5, Appli
18	539	13.2	1395	10	US-09-974-300-907	Sequence 907, App
19	534	13.1	2127	12	US-10-310-154-354	Sequence 354, App
20	533	13.1	1386	12	US-10-369-493-47058	Sequence 47058, A
21	505	12.4	1422	12	US-10-369-493-46784	Sequence 46784, A
22	496	12.2	1853	14	US-10-051-902-23	Sequence 23, Appl
23	496	12.2	1853	14	US-10-051-909-23	Sequence 23, Appl
24	494.5	12.1	2017	14	US-10-051-902-21	Sequence 21, Appl
25	494.5	12.1	2017	14	US-10-051-909-21	Sequence 21, Appl
26	493	12.1	1482	10	US-09-938-842A-795	Sequence 795, App
27	491.5	12.1	1914	14	US-10-051-902-19	Sequence 19, Appl
28	491.5	12.1	1914	14	US-10-051-909-19	Sequence 19, Appl
29	489	12.0	1566	10	US-09-938-842A-1769	Sequence 1769, Ap
30	489	12.0	1872	14	US-10-051-902-27	Sequence 27, Appl
31	489	12.0	1872	14	US-10-051-909-27	Sequence 27, Appl
32	486.5	11.9	1449	12	US-10-369-493-41339	Sequence 41339, A
33	475	11.7	1644	10	US-09-938-842A-320	Sequence 320, App
34	473.5	11.6	1395	12	US-10-369-493-46991	Sequence 46991, A
35	469.5	11.5	2051	12	US-10-310-154-356	Sequence 356, App
36	465	11.4	1260	12	US-10-369-493-46835	Sequence 46835, A
37	465	11.4	1931	12	US-10-310-154-355	Sequence 355, App
38	465	11.4	2063	14	US-10-051-909-33	Sequence 33, Appl
39	463.5	11.4	2089	14	US-10-051-902-25	Sequence 25, Appl
40	463.5	11.4	2089	14	US-10-051-909-25	Sequence 25, Appl
41	462	11.3	1839	12	US-10-369-493-25706	Sequence 25706, A
42	459	11.3	285	9	US-09-923-876-6145	Sequence 6145, Ap
43	459	11.3	285	12	US-09-923-876-6145	Sequence 6145, Ap
44	453.5	11.1	1905	9	US-09-734-569-171	Sequence 171, App
45	451	11.1	1377	12	US-10-369-493-39348	Sequence 39348, A

ALIGNMENTS

RESULT 1
US-10-051-909-31
; Sequence 31, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 2777
; TYPE: DNA
; ORGANISM: Zea mays
US-10-051-909-31

Alignment Scores:

Pred. No.:	0	Length:	2777
Score:	4075.00	Matches:	800
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-051-909-32 (1-800) x US-10-051-909-31 (1-2777)

```
QY      1  ILeaRgSerGIySerTPrLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
      12  ATTCGAGCGGCTCTTGCTGCGATCCAAACGCCCTTCACCCCTGATCTGACCGGAGG 71
QY      21  GlnArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
      72  GAGCGGCTCCTCCGTCAGTGTCTTGCTTGCCCTGGGCTCTTCCGCTGCTGCTGT 131
QY      41  SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
      132  TCTTCACAGAGACCGCGTGACCTCGACGATATCTT3GAGGACAGAGATGCGGGGCTGT 191
QY      61  LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
      192  CTTCGCGCATAGTCGCTCCATCGGCAATCTAT3CAGGGGTGGGACAAATGCCACGATC 251
QY      81  AlaAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
      252  GCACTGCTGTTCTGTATATAAAGAAGAAATTCATTCATTGCAAAATGAGCCCACTGTGAG 311
QY      101  GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro 120
      312  GGACTAATTGTGTCAATGTCAATTCGGCGCCACCATCGTTACTACATTCCTCCGGGCA 371
QY      121  LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
      372  TTAATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTTCAATTCCTGACTTCTTC 431
QY      141  SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheVal 160
      432  AGCGGCTCATCATGCTATAGGTCTCTAATGCTATGCTGCTGCTGTGGCAGCTTGTA 491
QY      161  AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
      492  GATGGATTGGTATTTGGCTTGCTGCTCACGCTTGCTGCTTGTACATTTCAGAAATAGCC 551
QY      181  ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMet 200
      552  CCTTCGAGATTAGAGGTTTGCTGAATACTACCAACAATTCAGTGATCAGGAGAAATG 611
QY      201  PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220
      612  TTCTTGTCATACTGCATGGTGTTGGATGTGCCCTGTGCCATCACCCGATTGGAGATT 671
QY      221  MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
      672  ATGCTTGCTGCTCGCGAATCCTTCATTGTTCTTGGTTTGAACAATATTTTACTT 731
QY      241  ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260
      732  CCTGAATCTCCAAGATGGCTCGTTAGCAAAAGTCCGATGCGCAGAGCAAAAAAGGTGTG 791
QY      261  GlnLysLeuArgGlyLysAspAspValSerGlyLysLeuSerLeuLeuLeuGluGlyLeu 280
      792  CAATAAGTTACGGGGGAAGACGATGTCTCAGGTGAATGTCTTCTTCGAGGGGTG 851
QY      281  GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
      852  GAGGTGAGAGACACTTCATTTGAAGAGTACATCATTTGGAAGCTGCCACCGAGGAGCC 911
QY      301  AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
      912  GATGATCTTGTACTGACGGGTGATAAGGAACAATCACACTTATGGGCTGAAGAGGC 971
      Db
```

```
QY      321  GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
      972  CAGTCATGATTGCTCGACCTTCTTAAGGAGCCATCATGCTTGGAAGTGTGCTTCTCTT 1031
QY      341  AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
      1032  GCATCTGTCATGGAGCATGTGTGAACAGAGTGTACCCCTTATGATTCGATTGTGACA 1091
QY      361  LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
      1092  CTTTGTGATGTTCATGAGAATATGCTCAAGCTGAGAGAGTATGAGACACATTG 1151
QY      381  PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
      1152  TTTCAAACTTTGGAAGTATGTTCAAGTGCACAGATCAGCATGCCAAATAAGACAGTGG 1211
QY      401  AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyValGlyGlyAsp 420
      1212  GATGAAGAGAAATCTTCATAGGATGACGAGAGTACCGCATCTGATGGTGCAGAGGTGAC 1271
QY      421  TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
      1272  TATGAGACAAATCTTCATAGCCCATGTGCTGTCCAGGCGACGACAGAGTGGGAAG 1331
QY      441  AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
      1332  GACATTGTGACCAATGTGCACCCGTGAAGTGTGAGCATGAGAAGGCAAAACCTCTTA 1391
QY      461  GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrProGlnLeuAlaTrp 480
      1392  GGGAGGAGTGAGATGGTGTGAGCAGACATGATCCGTTGGGAGTGGACCTTGCTGG 1451
QY      481  LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
      1452  AAATGTCAGAGAGAAGAGGTGAGAAATGTTGAAGAAGAGTGTTCAAAAGAGTCTAC 1511
QY      501  LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
      1512  TTGCACCAAGAGGAGTTCTTGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGT 1571
QY      521  GlyAspValPheGluGlySerGluPheValHisAlaAlaValSerGlnSerAla 540
      1572  GCGGATGTTTGAAGGTAGTAGATTGTACATGCTGCTGCTTGAAGTCAAGTCAAGCA 1631
QY      541  LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
      1632  CTTTCTCAAGGCTTGCTGTAACACGAGTGTCAAGTGTCCATGCTCACCCATCT 1691
QY      561  GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla 580
      1692  GAGTAGCTGCCAAAGGTTCAAGTTGAAGAATTTGTTGAACCTGAGTGAAGCGGTGCC 1751
QY      581  LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600
      1752  CTGTTAGTCGGTGTGGAATTCAAGATCCTTCAACAGTTTGCTGGAATAAACGGGTCTCG 1811
QY      601  TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
      1812  TACTATAACCCCAAAATCTTGAGCAAGCTGTGTGGCAGTTATTTCTTCCAAATTTGGT 1871
QY      621  LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
      1872  CTCAGCTCGGCATCAGCATCATCTTGATCAGTTCTCTCAGTACCTTAATGCTTCTCT 1931
QY      641  CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
      1932  TGCATTGGCTTGCCATGCTGCTATGATCTTCCGGAAGAAGTTTTGCTGCTAGGC 1991
QY      661  ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
      1992  ACAATTCCAATCTTGATAGCATCTCTAGTTATCTCTGTTGTGTCCAAATCTAATTGATT 2051
      Db
```


QY 681 GLYThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 2052 GGTACACTAGCCCATGCTTGTCTCTCCACCGCTCAGTGTATGCTACTTCTGCTTC 2111
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2112 GTTATGGATTGGTCCCATCCCAACATTTATGTCAGAGATCTTCCACCAAGGTT 2171
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 2172 CGTGGCTCTGTATGCCATTGTGCTTACATCTGGATCGAGATATCATCTCACC 2231
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 2232 TACAGCCTTCCTGTATGCTGAATGCTATTGAGTGGCGGCTTTTCAGCATATATGCA 2291
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2292 GTCGTATGCTGATTCTTCTTGTGTCTTCTTAAGGTCCCTGAGACCAAGGGATG 2351
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGluAlaAlaLysAla 800
Db 2352 CCCCTTGAGTTATTCAGCAATTCTTTCAGTTGGTGGCAAGCAAGCGGCTGCAAAAGCC 2411

RESULT 2
US-10-310-154-350

; Sequence 350, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shishieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luechy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Phillip W.
; APPLICANT: Padmavathi, Manthikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan

; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 350.
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (331)..(2565)
; OTHER INFORMATION:
US-10-310-154-350

1024 Feb

Alignment Scores:
Pred. No.: 0 Length: 2908
Score: 4059.00 Matches: 797
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 12 Gaps: 0

US-10-051-909-32 (1-800) x US-10-310-154-350 (1-2908)

QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
Db 166 ATTCGAGCGGCTCTTGGCTTGACAGTCCAGACGCCCTTCAACCCCTGATCTGACCGGAGG 225
QY 21 GluArgLeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCys 40
Db 226 GAGCGGCTCTCCGTGAGTGTCTTGTGCTTGGCGGCTTCCGCTGCTGTGT 285
QY 41 SerSerGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
Db 286 TCTTACAGAGCGCGGTGACCTCGACGATATCTTGAGAGACAGATGTCGGGCTGTT 345
QY 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIle 80
Db 346 CTGTGCGCATAGTCGCTCCATCGGCAATCTATGACAGGGGTGGACATGCCACCATC 405
QY 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
Db 406 GCAGCTGCTGTCTGTATATAAAGAGAAATTTCAATTGCAAAATGAGCCCATGTGAG 465
QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro 120
Db 466 GGAATAATTGTCAATGTCACTTATCGGCCACCATCGTTACTACATCTCCGGGCA 525
QY 121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
Db 526 TTATCAGACTCGATTGGCCGACGCGCTATGTTATCTCTTCAATTCTGACTTCTTC 585
QY 141 SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheVal 160
Db 586 AGCGGCTCATCATGTATGCTCTCTAATGTATGTCTGCTGTGGCAGCTTCGTA 645
QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 646 GATGATTGTGATTGGCTTGGCTGTCAAGCTTGTGCTTGTACATTTCAAAATAGCC 705
QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
Db 706 CTTGAGAGATTAGAGGTTTGTGTAATACACTACCAATTCAGTGAATCAGAGGAATG 765
QY 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220
Db 766 TTCTGTATACATGATGTGTGTTGGATGTCCTGTGCGCATCACCGAATGAGAATT 825
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240

Db 826 ATGCTTGCTGCTCGCGATACCTTCATTTCTTCTTGGTTTGACAATATTTTATCTT 885
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260
Db 886 CTGAATCTCCAAGATGGCTGTTAGCAAAAGGTGCGATGGCAGAGGCAAAAAGGTGTG 945
QY 261 GlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
Db 946 CAAAAGTTACGGGGGAAAGACGATGTCAGGTGAATGTCCCTTCTTCGAAGGTTG 1005
QY 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 1006 GAGCTTGAGAGACACTTCCATTTGAAGATACATCATTTGACCTGCCACCGAGGAGCC 1065
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
Db 1066 GATGATCTTGTTACTGACGGTGATAAGAACAAATCACACTTTATGGGCTGAAGAGGC 1125
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 1126 CAGTCATGGAATTGCTCGACCTTCCAAAGGACCAGCATGCTTGGAAGTGCTTCTCTT 1185
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 1186 GCATCTCGTCATGGAGCATGGTGAACAGAGTGTACCCCTTATGTGATCCGATGTGACA 1245
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
Db 1246 CTTTGTGTAGTGTCCATGAGAATATGCTCAAGCTGAGGAAGTATGAGAGACACATG 1305
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
Db 1306 TTTCCAAACTTTGGAAGTATGTCAGTGTCAAGATCAGCATGCCAAATAGAGCAGTGG 1365
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyLysAsp 420
Db 1366 GATGAAGAAGATCTTCATAGGATGACGAGAGTACCGCATCTGATGCTGACAGAGGTGAC 1425
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
Db 1426 TATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGG3CAGGCCAACAGGTGCGGAAGG 1485
QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
Db 1486 GACATTTGTGCACCATGTCACCCGTGAAGTGTCTTTAGCATGAGAAAGGCCAAGCCTCTTA 1545
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrp 480
Db 1546 GGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGTGGGGGATGGCAGCTTGCTGG 1605
QY 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 1606 AAATGTCAGAGAAGAGAGTGAATGTAGAAAGAGAGGTGTTCAAAAGAGTCTAC 1665
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 1666 TTGCAACCAAGAGGAGTTCCTGCTCAAGAGGGGCTCAATTTGTTCACTTCCCGGTGCT 1725
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAla 540
Db 1726 GCGGATGTTCTTGAGGGTAGTGAAGTTGTACATGCTGCTGTTAAGTACAGTACGA 1785
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 1786 CTTTCTCAAAAGGCTCTGTGTAACACGCAATGTCAGATGCTGCATGTTCAACCACTCT 1845
QY 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla 580
Db 1846 GAGGTAGCTGCCAAAGCTTCACGTTGGAAGAATTGTTTGAACCTGAGAGTGAAGCGGTGC 1905
QY 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600

Db 1906 CTGTTAGTCGGTGTGGAATTACAGATCCTTCAACAGTTTGCTGGAATAAACGCTGTTCTG 1965
QY 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
Db 1966 TACTATACCCACAATCTTTGAGCAAGCTGGTGTGGCACTTATCTTCCAAATTTGGT 2025
QY 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
Db 2026 CTCAGCTCGGCATGACATCCATCTTGATGACGTTCTCTCACTACCTTACTAATGCTTCCT 2085
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Db 2086 TGCATTGGCTTGGCATGCTGCTTATGATCTTTCCGGAAGAGTTTTCCTGCTAGGC 2145
QY 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
Db 2146 ACAATTCCAATCTGATAGCATCTCTAGTTATCTGTTGTGTCCAATCTAATGATTTG 2205
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 2206 GGTACACTAGCCCATGCTTGTGCTCTCCACCGTCAGTGTATCGTCTACTTCTGCTTCC 2265
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2266 GTATGAGATTGGTCCCATCCCAACATTTATGTGCAAGAGATCTTCCAACAGAGGT 2325
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 2326 CGTGCCCTGTGATTTGCCATTTGTGCTTACATTTCTGGATCGAGATATCATCTGTACC 2385
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 2386 TACAGCCTTCTGTGATGCTGAATGCTATGTAGACTGGCGGGTGTTTTACAGCATATATGCA 2445
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2446 GTCGTATGCTGATTTCTTGTGTTGTTGTTCTTAAAGTCCCTGAGACAAAGGGGATG 2505
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2506 CCCCTGAGGTATATACCGAATTCTTTGAGTTGGTGGAGCAAGCGGCTGCCAAAAGCC 2565

RESULT 3
US-10-051-902-1
; Sequence 1, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure
; LOCATION: (638)
; NAME/KEY: unsure

LOCATION: (669)
NAME/KEY: unsure
LOCATION: (771)
NAME/KEY: unsure
LOCATION: (822)
NAME/KEY: unsure
LOCATION: (856)
NAME/KEY: unsure
LOCATION: (889)
NAME/KEY: unsure
LOCATION: (896)
NAME/KEY: unsure
LOCATION: (944)
US-10-051-902-1

Alignment Scores:

Pred. No.:	0	Length:	2824
Score:	3517.00	Matches:	684
Percent Similarity:	95.18%	Conservative:	27
Best Local Similarity:	91.57%	Mismatches:	34
Query Match:	86.31%	Indels:	2
DB:	14	Gaps:	2

US-10-051-909-32 (1-800) x US-10-051-902-1 (1-2824)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db ATGGGGGGCGCGGTGATGCTCCCATCGCGGCTCTATCGCAACTGTGTCAGGGCTGG 297
QY 76 AspaSnAlaThrIleAlaAlaValLeuTyrIleYslySgluPheGlnLeuGlnAsn 95
Db GACAAATGCGAATGCTGAGCCGCTCTGTACATAAGAAAGAAATTCACCTGCGAGAC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db GAGCCTCTGATCGAAGGCTCATCGCCATGTTCCATTTGGGGCAACAGTCATACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db ACATCTCCGGGGCCAGGGCTGACTGCGTTGGTAGGAGGCCCATGCTGTCGCTCGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db GTCTCTACTTCTGTCAGTGGCTGTGATGCTTTGGGCGCCAATTGTACATCTTGTCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db CTGCAAGGCTCATGATGGTTCGTATCGGTTGGCGGTCACACTTGTCTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db ATCTCCGAACACTGCACCGCAGANATTTCTGGGCTGNTNGAACAAGTTCGCCGCTTC 657
QY 195 SerGly--SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db ATGGGGGTCAAGNGAGGGATGTTCTCTCTACTGCAAGTGTGGATGTCCCTCATG 717
QY 214 ProSerProAspTyrPargIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db CCCAAACCTGATGGAGGCTCAAGTCTGTGAGTTCGTGATCCCGTCACTTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrPheValSerIysGlyArgMet 253
Db GGAAGTGAAGTCTTCTACTTCTGAAATCACCAAGGTGGCTGTNAGCAAGAGAGATG 837
QY 254 AlaGluAlaIleValLeuGlnIleValArgGlyLysAspAspValSerGlyGlnLeu 273
Db GCGAGGCGAAGAGAGAGTGTGCAAGGCTGCGGGAAGAGAGATGTCTCANGGAGANG 897
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyLysPheThrSerIleGluGluTyrIle 293
Db GCTCTTCTAGTTGAAGGTTGGGGGTCCGTAAGATACAGTATTTNAGAGTACATCATTT 957

QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db GGAAGTGAAGTCTTCTACTTCTGAAATGACCGGTGATGAAGCAAAATCAC 1017
QY 314 LeuTyrGlyProGluGluGlyIleSerTyrIleAlaArgProSerIysGlyProIleMet 333
Db CTTATGGGCTGAAGAGAGCCAGTCATGATGCTCGACCTTCTTAAGGAGCCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db CTTGAGAGTGTGCTTCTCTGATCTCGTCATGGAGCATGTGAACCAAGTGTACCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db CTTATGATCCGATGTGACACTTTTGTGATGTTCATGAGATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db GGAAGTATGAGAGCATGTTGTTCCAACTTTGGAAGTATGTTGAGTGTCAAGATCAG 1257
QY 394 HisAlaLysAsnGluGlnTyrAspGluGluAsnLeuHisArgAspAspGluTyrAla 413
Db CATGCCAAAATGAGCAGTGGGATGAAGAAATCTTCATAGGATGACGAGAGTACGCA 1317
QY 414 SerAspGlyAlaGlyLysPyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db TCTGATGTGTCAGAGGTGACTATGAGACAAATCTCCATAGCCATTTGCTGTCCAGGCA 1377
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
Db GCAACAGTGCAGAGGAGGAGACATTTGTCACCATGTCACCGTGAAGTCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly 473
Db ATGAGAAAGCAAGCTCTTAGGGAGGATGAGATGTTGTGAGCAGACACTGATTCGGT 1497
QY 474 GlyGlyTyrPglLeuAlaTyrLysTyrSerGluLysGluGluAsnGlyArgLysGlu 493
Db GGGGATGCGAGCTTGTGGAATGTCAGAGAAGAGTGAAGATGTTGAAGAAAGAA 1557
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgLysSer 513
Db GGTGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGCTCAGAGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyLysPheValPheGluGlySerGluPheValHisAlaAla 533
Db ATTGTTTCACTTCCCGGTGTGCGCATGTTCTTGAAGGTAGTAGTTGTATGCTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db GCTTAGTAAGTACAGTCAACATTTTCTCAAGGCTTGTGTAACCAACGATGTCAAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe 573
Db GTCGCATGTGTTCAACCATCTGAGGTAGCTGCCAAAGTTTACGTTGAAGAATTGTTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db GAACCTGAGTGAAGGCTGCTGTTAGTGGGTGGAATTCAAGATCCTTCAACAGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db GCTGAATTAACGGTGTCTGTACTATACCCCAAAATTTTGAAGCAAGCTGTGTGCA 1917
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db GTTATTTCTTCCAAATTTGCTCTCAGCTCGCATCAGCATCATCTTGAATCATCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db ACTACCTTACTAATGCTTCTTGCATTTGCTTGCATGCTGCTTATGATCTTCCGGA 2037
QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673

|||||
Db 2038 AGAAGTTTGTGCTAGGCAATCCAACTCTGATAGACTCTAGTATCTGTT 2097
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAATCTAATGATTTGGGTACACTAGCCCAAGCTTTGCTCTCCACCATCACTGTT 2157
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATGCTACTTCTGCTGCTTGTGTTATGGGATTGGTCCCATCCCAACATTTATGTGCA 2217
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTCCAAACCAAGGTTGCGCTCTGTATGCGCATTTGCTTACATCTGG 2277
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATGGAGATATCATGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTGAGCTGGC 2337
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTCAGCATATATGCACTGCTATGCTGATTTCCTTTGTGTTGCTCTTAAAG 2397
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAAGGGATGCCCTTGAGGTATTACCGAATCTTTGACAGTTGTGCG 2457
QY 794 LysGlnAlaAlaAlaLysAla 800
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478

RESULT 4
US-10-051-909-1

; Sequence 1, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR APPLICATION DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure
; LOCATION: (638)
; NAME/KEY: unsure
; LOCATION: (669)
; NAME/KEY: unsure
; LOCATION: (771)
; NAME/KEY: unsure
; LOCATION: (822)
; NAME/KEY: unsure
; LOCATION: (856)
; NAME/KEY: unsure
; LOCATION: (889)
; NAME/KEY: unsure

; LOCATION: (896)
; NAME/KEY: unsure
; LOCATION: (944)
US-10-051-909-1

Alignment Scores:
Score. No.: 0 Length: 2824
Percent Similarity: 3517.00 Matches: 684
Best Local Similarity: 95.18% Conservative: 27
Query Match: 91.57% Mismatches: 34
DB: 86.31% Indels: 2
Gaps: 2

US-10-051-909-32 (1-800) x US-10-051-909-1 (1-2824)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 238 ATGGGGGGCGCGGTGATGTCGCCCATCGCGGCTCTATCGCACTTGCGAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 298 GACAATGCCACAAATTGCTGAGCCGCTCTGTATCAATAAAGAAGAAATTCACTGCAGAGC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTGTGATCGAAGGCTCATGCTGCCATGTTCTCATTTGGGGCAACAGTCATCACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCCGGGGCCCAAGGCTGACTGCTGTGTGAAGAGGCCCATGCTGTCGCTCGGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
Db 478 GTCCCTACTTCGTCAAGTGGGCTGTGATGCTTTGGCGCCAATTGTGTACATCTGCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 538 CTCGCAAGGCTCATTTGATGGGTTCCGATCGGTTTGGCGGTACACACTGTCTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAACCTGACCCGACAGANATTCTTGGGGCTGNTNGAACACGTTGCCGACGTT 657
QY 195 SerGly--SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATGGGGTCAAGAGGAGGATGTTCTCTCTCACTGATGATGTTGGAGTGTCCCTCATG 717
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 718 CCCAAACCTGATGAGGCTCATGCTTGAGTCTGTGCATCCCGTCACTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db 778 GGACTGACTGTCTTCTACTTGCCCTGAATCACCAAGGTGCTGTNAGCAAGAAGATG 837
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeu 273
Db 838 GCGGAGGCGGAAGAGAGTNTGCCAAAGGCTGCGGGGAAGAGAAGATGTCTCANGGGAGANG 897
QY 274 SerLeuLeuLeuGlnGlyLeuGluValGlyAspThrSerIleGlnGluTyrIleIle 293
Db 898 GCTCTTCTAGTTGAAGGTTTGGGGTCCGTAAAGATACACGTATTNAGGTACATCAT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlnGlnIleThr 313
Db 958 GGACCTGCCACCGAGGCAAGCCGATGATCTTTAACTGACGCGTGAATGAAGCAAAATCACA 1017
QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 1018 CTTTATGGGCGCTGAAGAAGCCAGTCATGATGCTCGACCTTCTAAGGACCCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValaGlnSerValPro 353

Db 1078 CTTGGAAGTGTCTTCTTCATCTCGTCATGGAGCATGGTGAACCAAGAGTGTACC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGATCCGATTGTGACACTTTTGTGAGTGTCCATGAGATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPhcGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGTATGAGAGCACAATTGTTCCAAACTTTGAGTATGTTCAAGTGTCAAGATCAG 1257
QY 394 HisAlaIysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluTyrAla 413
Db 1258 CATGCCAAAATGAGCAGTGGGATGAGAGAATCTTCATAGGATGACGAGAGTACGCA 1317
QY 414 SerAspGlyAlaGlyIysAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGTCAGAGGTGACTATGAGGACAAATCTCCATAGCCCATTTGCTCCAGGGCAG 1377
QY 434 AlaThrGlyAlaGluGlyIysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGCGAAGGGAAGACATTTGTGACCAATGTCACCGTGGAAGTCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyIysAspGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAAGGCAAGCCCTCTTAGGGAGGGTGAAGATGCTGTGAGCAGCAGCATATCGGT 1497
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGlnLysGluGlyIysAsnGlyArgLysGlu 493
Db 1498 GGGGGATGGCAGCTGTCTGGAATGCTCAGAGAAGAGGTGAGAATGCTGGAAGAA 1557
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
Db 1558 GGTGGTTCAAAAGAGTCTACTTGACCAAGAGGAGTCTTGCTCAAGAGGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1618 ATGTGTTCACTTCCCGGTGGTGGCATGTTCTTGAGGTAGTGAAGTTGTACATCTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTAGTAGTCAAGTCAAGCACTTTCTCAAGGGTCTTGCTGAACCAAGCATGTACAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaIleValSerArgTrpLysAspLeuPhe 573
Db 1738 GCTGCCATGTTCAACCACTCTGAGTAGTCTGCCAAGGTTCAAGTTGAAGAATTTGTTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db 1798 GAACCTGAGTGAAGGCGTCCCTGTTAGTGGGTGGAAATTCAGATCCTTCAACAGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1858 GCTGGAATAAACGGTGTCTGTAATAATACCCCAAAATTTCTTGAGCAAGCTGTGTGCA 1917
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCTTGTATCAGTTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACCTTAATAATGCTCTGATGCTTGGCTTGCATGCTGCTATGATCTTCCGGA 2037
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerIleValIleLeuVal 673
Db 2038 AGAAGGTTTTGCTGCTAGGCAACAATTCCAATCTTGATAGCATCTCTAGTTATCTGTT 2097
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAATCTAATTGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATGCTACTTCTGCTGCTTGTATGGGATTGTGTCATCCCAACATTTTATGTGCA 2217

QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTTCACCAAGGTTGCTGGCCCTGTATGTCATTTGTGCTTACATTTCTG 2277
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGAGATATCATGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTGGAAGTGGC 2337
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTACAGATATATGACAGTGTGATGCTGATTTCTTGTGTGCTTCTTAAAG 2397
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAGGGATGCCCCCTGAGGTATTACCGAATCTTTCAGATTGTTGCG 2457
QY 794 LysGlnAlaAlaAlaLysAla 800
Db 2458 AAGCAAGCGCTGCAAAAGCC 2478

RESULT 5

US-10-051-902-7
; Sequence 7, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-051-902-7

Alignment Scores:

Pred. No.: 5,53e-274 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 14 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-902-7 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 175 ATGAAGGTGCCGCTTGTGCTATGTGCCGCTTCCATTGTAATTTCTCCAAAGATGG 234
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 235 GATAATGCTACCATCGCCGGGCTAATGTTAATTAAGAAGACCTTGCTTTGGGAACA 294
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGGAAGGCTTGTGTTGGGCGCATGTCCTGATTGAGCAACGGTAATCAC 348
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSerSer 135
Db 349 ACATGCTCTGGCTATATAGCGATTGGCTGGCGGACCAATGATGATATCTCATCT 408
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
Db 409 GTGCTTATTTCTTGGGTGTTGTTGATGCTGTGTCGCCCAATGTGTATGTGTGTC 468

```
QY 156 LeuAlaargPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 469 TTGGCGAGGCTACTGATGATTTGGGATGGCCCTGTGTGACTCTTGTCCCGGTCTAT 528
QY 176 IleSerGluIleAlaProSerGluIleargGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAACACGGCCGCTGAATAAGGGGCTGTGAATACGCTTCTCAGTTCACT 588
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 589 GGCCTGGAAGGAATGTTTGTGCTACTGTATGTTTGGCATGTCTATTGAGTCCCGCG 648
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 649 CTTAGCTGAGGCTCATGTGGGTTCTGTCTATCTCTCTCTCTCTTGTATTGTCATG 708
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMetAlaGlu 255
Db 709 ACCATTTTCTTCTGCGGAGTCTCTCGGTGGCTGTGTAGCAAGAGGATGCTCGAG 768
QY 256 AlaIleValValLeuGlnIleValSerGlyIleValAspAspValSerGlyIleLeuSerIleu 275
Db 769 GCTAAGAAAGGTGCTCCAAAGATTCGCCGGAAGGAGATGTCTCAGCGGAGATGGCATTG 828
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db 829 CTGGTTGAAGGTCTCGGATTTGGGGTGATACATCTATCGAAGATACATAATTGGCCCT 888
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIleGluGlnIleThrLeuTyr 315
Db 889 GCTGACGATGTGGCTGATGTGATGAACATGCACACAGAAAGATAAATTCGATTATAT 948
QY 316 GlyProGluGluGlyIleSerTrpIleAlaArgProSerIleGlyProIleMetLeuGly 335
Db 949 GGATCCCAAGCAGGCTTCTTGGTTATCAAAACCTGTCTAGTGAAGAGTCTATGGC 1008
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db 1057 GATCTCTGTGACACTGTGTGATGATTCATGAGAAGCTCCCGAGACAGAGCAGA 1116
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTGGAAGCATGTTTCAGCACTGTGAGCCG 1176
QY 394 HisAlaIleAsnGluIleTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1177 CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAGGGAAGTGAAGACTACATG 1236
QY 414 SerAspGlyAlaGlyIleAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1237 TCAGATGCAACCCGCTGGGACCTCGATGATAATTGACAGCTCTTAATCTCAGCCAA 1296
QY 434 AlaThrGlyAlaGlyIleValAspIleValHisIleGlyHisArgGlySerAlaLeu--- 452
Db 1297 ACAACAAAGCCTTGA---AAAGACTTACCTCTCTCTCTCTCTCTCTCTCTCTG 1353
QY 453 SerMetArgArgGlnThrLeuLeuGlyGlyGly---GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGCTCAGGTGAGCAAGGTGTAGTACAGGT 1413
QY 472 IleGlyGlyIleTyrPheGlnLeuAlaTrpIleTyrSerGluIleGlyGluAsnGlyArg 491
Db 1414 ATTGCTGTGTGGCTGCAACTGCAATGAATGAGTGAATAA---GGTGAAGATGAAAA 1470
QY 492 LysGluGlyGlyPheIleArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAAGAGGTTTAAAGATTTATTTACATGAGAGGAGTTTCTGCATCTCTGCT 1530
```

```
QY 512 GlySerIleValSerLeuProGlyGlyIleAspValPheGluGlySerGluPheValHis 531
Db 1531 GGATCCATTGTATCGATTCGCCGTGAAGC-----GAATTTGTCCAG 1572
QY 532 AlaAlaIleValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMet 551
Db 1573 GCTGCTGCCCTTGGTGAAGCAACCCGCTCTTACTCCAAAGAGCTTATGTAGACACCCA 1632
QY 552 SerAspAlaIleMetValHisProSerGluValAlaAlaIleGlySerArgTrpIleAsp 571
Db 1633 GTTGGGCTGCAATGTGTCAACCATCTGAGACAGCTTCAAGGGGCCAAGTTGGAAGCT 1692
QY 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTCTTGAACCAAGGCTTAAGCATGATGATGTTGTTGAGTTGGAATACAAATACTTCAG 1752
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly 611
Db 1753 CAGTTTCAGGATTAATGGGTTCTATATTAACACCTCAAAATCTTGAAGAGCCGGT 1812
QY 612 ValAlaValIleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTCTTCTTCAATATAGCATGATGGCTCAGAGTCGCATCATCTTATCAGT 1872
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeu 651
Db 1873 GCTTACACACCTTCTGATGCTTCCCTGTATAGCGGTAGCCATGAGCTCATGATGTT 1932
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAAGAGGCACTGCTACTTACTACAAATCCCGCTGATGTGTCTACTCATTTAT 1992
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGCTCATTTGAAGCTGTGTAATTTGGCAATGTGCCCATGCAGCAATCTCAACAGTA 2052
QY 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TGCCTTGTGTTATTTCTGCTGTGTGATGGGTATGACCAATTCAAACATCCTT 2112
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTCCCACTAGAGGTGCGGCTCTGCATGTCTATCTGTGATTAAGT 2172
QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGATGAGACATCATCATCATCACTACTGCTGCTGATGCTGCGCTCTTAAGA 2232
QY 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db 2233 CTGTGTGTGTATTCGCCATTTACGCAAGTTGTTGTTTCATCTCTGTGATATTGTGTT 2292
QY 772 LeuIleValProGluThrIleGlyMetProLeuGluValIleThrGluPheAlaVal 791
Db 2293 TTGAAGTTCCAGAAACAAAGGCGATGCCCTTGAAGTCATCTTGAAATCTTTCTGTT 2352
QY 792 GlyAlaIleGlnAlaAla 798
Db 2353 GGAGCAAGCAGCTGCTTCT 2373
```

RESULT 6

```
US-10-051-909-7
; Sequence 7, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051, 909
```

; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-909-7

Alignment Scores:
Pred. No.: 5.53e-274 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 14 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-909-7 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 175 ATGAAAGGTGCCGCTGTTGCTATTGCCGCTCCATTGTAATTCTCCAGGATGG 234
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 235 GATTAATGCTACCATCGCGGGCTAATGTTACATTAAAGAACCTTGCTTGGGACA 294
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGGAAGGCTTGTTGGGCATGTCCTGATGGAGCAACGGTAATCAC 348
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 349 ACATGCTCTGCTCTATAGCGGATGGCTCGGCGGACCATGATGATATCTCATCT 408
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPserProAsnValTyrValLeuLeu 155
Db 409 GTGCTCTATTCTTGGGTGGTTGGTGATGCTGTGTCCTCCCAATGTATGTGTGTC 468
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 469 TTGGCGAGGCTACTGATGATGGATTGGGATGGCCTTGCTGACTGACTCTGCCGCTAT 528
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAACGCGCGCTGTAATAAGGGGCTGTTGAATACGCTTCCAGTTCACT 588
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 589 GGCCTCTGAGGAATGTTTGTCTACTGATGTTTGGCATGTCATTGAGTCCCGCG 648
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 649 CTTAGCTGAGGCTCATGCTTGGGCTTCTGTCTATTCTCTCTCTGTTGTCATTG 708
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 709 ACCATTTTTCTTCTGCCGAGTCTCTCGGTGGCTGTGTCAGCAAGGAGATGCTGAG 768
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
Db 769 GCTAAGAGGTGCTCCAAAGATTGGCGGAGGAGGATGTGTCAAGCGAGATGGCATTG 828
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGlnGluTyrIleIleGlyPro 295
Db 829 CTGGTTGAAGCTCTCGGATTTGGGGTGATACATCTATCGAAGATACATAATTGGCCCT 888
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlnGlnIleThrLeuTyr 315
Db 889 GCTGACGATGTGGCTGATGTCATGACATGCAACAGAGAAAGATAAATTGATTATAT 948

QY 316 GlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335.
Db 949 GGATCCCCAAGCAGGCTTCTTGGTTATCAAACTGTCACTGACAGAGTTCTATTGGC 1008
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 1009 -----CTTGGCTCACACCATGAGCATCATCAACCAAGCATGCCCCCTCATG 1056
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db 1057 GATCCTCTGTGACACTGTTTGGTAGCATTCATGAGAAGCTCCCCGAGACAGACAAGA 1116
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1117 GGAAGCATGCGAAGCACTGTGTTCCAAATTTTGGAAAGCATGTTCAAGCACTGTGAGCCG 1176
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1177 CATGCTAAATTTGAACAATGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG 1236
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1237 TCAGATGCAACCCGTGGGAGCTCCGATGATTAATTGCACAGTCTTTAATCTCACGCCAA 1296
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeu--- 452
Db 1297 ACAACAAGCCTGAA---AAAGACTTACCTCTCTCTCCCTCCATGGCAGTATCCTTGGC 1353
QY 453 SerMetArgArgGlnThrLeuLeuGluGlyGly---GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
QY 472 IleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGluAsnGlyArg 491
Db 1414 ATGTGGTGCTGGCACTGGCATGGAATGCACTGATAA---GGTAGGATGGAATAA 1470
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAAGAGGAGTTTAAAGATTATTATTACATGAGAGGAGATTCTGCATCTCGTCGT 1530
QY 512 GlySerIleValSerLeuProGlyGlyLysPValPheGluGlySerGluPheValHis 531
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGC-----GAATTGTCCAG 1572
QY 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db 1573 GCTGCTGCTTGGTGAAGCCAAACCCGCTCTTACTCCAAAGAGCTTATTGATGACACCCA 1632
QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db 1633 GTTGGCCCTGCAATGGTTCAACCATGTGAGACAGCTTCAAAAGGGGCCAAGTTGGAAGCT 1692
QY 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTTCTTGAACCAAGGCTTAAGCATGATGTTGTTGGAAGTTGAATACAAATTACTTCAAG 1752
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGly 611
Db 1753 CAGTTTCAAGGATAAATGGGTTCTATATTACACACCTCAATCCTTGAAGAGCCGGT 1812
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTCTTCTTTCAGATATAAGCATTTGGCTCAGATCGCATCATCTTATCACT 1872
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGAGTGT 1932
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAAGAGGAGTGGTACTTACTACAAATCCCGGTGATGTTGTCTACTCATTTAT 1992
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691


```
Db      1993 TTGGTCATTGGAAGCCTGTAAATTTTGGCAATGTCGCCCATGCAACATCTCAACAGTA 2052
QY      692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db      2053 TGGCGTGTGGTTATTTCTGCTGCTTTGTGATGGGTATGAGCAATTCACAAACATCCTT 2112
QY      712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db      2113 TGCTCAGAGATTTCCCACTAGAGGTGCGTGCCCTTGCAATGCTATCTGTGCATTAGTG 2172
QY      732 PheTrrPIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db      2173 TTCTGATGAGACATCATCATCACTACTGCTGCTGATGCTCGGCTCTTACGA 2232
QY      752 LeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPhe 771
Db      2233 CTGTGTGTGTATTCGCCATTTACGCACTGTTGTTCACTCGTGATATTGTGTTT 2292
QY      772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaVal 791
Db      2293 TTGAAGGTTCCAGAACAAGGCGATGCCCTTGAAGTCATCTCTGAATTCTTTCTGTT 2352
QY      792 GlyAlaLysGlnAlaAlaAla 798
Db      2353 GGAGCAAGCAGGCTGCTTCT 2373
```

RESULT 7

```
US-09-938-842A-1315
; Sequence 1315, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1315
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1315
```

Alignment Scores:

```
Pred. No.: 1.61e-264 Length: 2190
Score: 2584.00 Matches: 512
Percent Similarity: 80.78% Conservative: 89
Best Local Similarity: 68.82% Mismatches: 123
Query Match: 63.41% Indels: 20
DB: 10 Gaps: 8
```

US-10-051-909-32 (1-800) x US-09-938-842A-1315 (1-2190)

```
QY      56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db      1 ATGAGTGGAGCTGTGCTTGTGCTATTGCTGCTGCTGTTGGCAACTTGTACAGGATGG 60
QY      76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db      61 GATPAACGCACTATTGACGAGAGCTGTGTGTACATMAAAGAGTTTAATTGAGAGT 120
QY      96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
```

```
Db      121 AATCCATCAGTGGAGGCTTAATTGTGGCGGATGTCACTTATTGTGCTACTCTGATTTACA 180
QY      116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db      181 ACATGCTCTGAGGGGTAGCTGATTTGGCTTGTCGCCCTCCATGATTAATATTGCTCTCA 240
QY      136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrrPserProAsnValTyrValLeuLeu 155
Db      241 ATTCTCTACTTGTGTGCTCTCTAGTAATAGCTAAGTCTCCGAATGTTATGTGTGCTC 300
QY      156 LeuAlaArgPheValAspGlyPheGlyTrrIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db      301 TTAGGAAGGTTGTAGATGATTGGGGTTGCTTGTCACACTGTTCCTATTAT 360
QY      176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db      361 ATATCTGAGACTGCACCACTGAGATTAGGGGACTGTGAATACGCTACCGCAGTTCACT 420
QY      196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db      421 GGCTCTGAGAGGATGTTCTTATCTTACTGTATGCTTTTCGAATGCTGTGATGCCATCA 480
QY      216 ProAspTrrPargIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db      481 CCTAGCTGAGATGATGCTTGTGCTGCTTTCATCCCTTCCCTGTCTTTTCTTCCTC 540
QY      236 ThrIlePheTyrLeuProGluSerProArgTrrLeuValSerLysGlyArgMetAlaGlu 255
Db      541 ACGGCTCTTCTTCTGCCCGAGTCCCAAGGTGGCTCGTGAGCAAAAGGTCAATGCTTGAA 600
QY      256 AlalysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyLysLeuSerLeu 275
Db      601 GCAAGCGGCTTCTCAGAGACTGCGGTGTCGCAAGATGTGTCTGTGATGCGCTTGG 660
QY      276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db      661 TTGTTGAGGGGTCTTGAATTGGAGGTGAACAACCATAGAGGAATATTAATTGCTCCC 720
QY      296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db      721 GCGATGAAGTACTGATGATCATGATATAGCTGTGATTAAGATCAATTAAATTATAT 780
QY      316 GlyProGluGluGlnSerTrrIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db      781 GGTGCAAGAGAGGCTGAGTTGGTGTGAGCCAGTCAAGAAGG-----GGA 828
QY      336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
Db      829 AGCACTATGAGCTTTTGTCTCGCCATGGAAGTACAATGAGCAGAGGCAAGCTCATTTG 888
QY      355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
Db      889 ATTGATCCTCTGTCCACACTGTTTGGAGCGTTCACGAGAAGATGCCGACACT---GGA 945
QY      375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
Db      946 AGCATGAGAGTGCCTTGTTCACCATTTTGGGAGTATGTTCAGTGTGGAGGGAATCAA 1005
QY      395 AlaLysAsnGluGlnTrrPaspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
Db      1006 CCAAGACATGAGATTGGAGTGAAGAGAACTTGTGGAGAAGGTGAGATTATCCATCC 1065
QY      415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db      1066 GAC---CATGAGATGATTCTGAAGATGATCTTCATCTCCGTGATCTCACGTCACAAAG 1122
QY      435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
Db      1123 ACAAGCATGAG---AAAGACATGCCCTCACACTGCTCATGAGACTCTTTCTACCTTCAGA 1179
QY      455 ArgArgGlnThrLeuLeuGlyGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
```

Db 1180 CATGAAGTCAAGTCAGGAGCTCAAGGGGAGAGCGGGTAGTATGGGATTGGAGGT 1239
QY 475 GLYTrpGlnLeuAlaTrpIysTrpSerGluIysGluIysGlnIysArgIysGluGly 494
Db 1240 GGATGGCAAGTGGCATGGAATGACGGAAAGAGAATGAATCGGACAGAAAGAA--- 1296
QY 495 GLYPheIysArgValTyrLeuHisGlnGluIysValProGlySerArgArgGlySerIle 514
Db 1297 -----GAAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329
QY 515 ValSerLeuProGlyGlyGlyAspValPheGluIysSerGluPheValHisAlaAlaAla 534
Db 1330 GTTTCATGCTGCTGGTGGATGGAACCGGTGAG--GCAGATTTTGTACAGACGCTGCT 1386
QY 535 LeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMetSerAspAla 554
Db 1387 TTGGTTAGCCAAACGAGCTCTTTATTCCAAAGACCTTCTCAAAAGACATACATTGGTCT 1446
QY 555 AlMetValHisProSerGluValAlaAlaIysGlySerArgTrpIysAspLeuPheGlu 574
Db 1447 GCTATGTTACATCCATCCGAA--ACAACTAAGGGGTCAATTGGCATGATCTTCATGAT 1503
QY 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAla 594
Db 1504 CCTGAGTCAAGCGTGCATTAAGTGTAGAGAGTGGACTTCAATACTTCAGCAGTTCTCA 1563
QY 595 GlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGlnIleAlaGlyValAlaVal 614
Db 1564 GGCATCAACGAGGAGTCTTACTACACACCGCAATCCTTGAGCAGCGGGTGTCCGGATC 1623
QY 615 IleLeuSerIysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
Db 1624 CTACTATCGAACATGGGAGTTAGTTCTTCTCAGCATCTTACTTATAAGTGCATTGACA 1683
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
Db 1684 ACCTTTGTATGTTACCTGCAATAGCTGTTCGAATGAGGCTCATGATCTTCTGTGCA 1743
QY 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
Db 1744 AGGACCTTGCTTCTCAACGACGATACCAATCCTGATAGCATCTTATGTTTAGTAATC 1803
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
Db 1804 TCAATCTGTGTACATGAACAGCATGTGTGACGCGGTCTTATCAACCGTAAGCGTTGTG 1863
QY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
Db 1864 CTCTACTTCTGCTTCTTGTGATGGGTTTCGGTCTCTGCTCCAAACATCCTCTGTTCAGAG 1923
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
Db 1924 ATTTTCCAACTCGAGTCCGCGGAATGTGCATCGCCATCGCCACTCGCACTCTTGATC 1983
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
Db 1984 TGTGACATTAATCGTCACTTACAGTCTCCCGTGTGCTCAATCCATTGGACTAGCTGCT 2043
QY 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuIysVal 774
Db 2044 GTGTTTGAATGTACGCAATCGTATGTGCATTTCAATGGGTCTTGTGTTCATTAAAGTC 2103
QY 775 ProGluThrIysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaIys 794
Db 2104 CCGGAAACTAAAGGCATGCCACTTGAAGTCAATCAGAGATTCTTTCTGTTGAGCTAGA 2163
QY 795 GlnAlaAlaAla 798
Db 2164 CAAGCTGAAGCT 2175

RESULT 8
US-09-938-842A-2254
; Sequence 2254, Application US/09938842A

; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2254
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-2254

Alignment Scores:
Pred. No.: 1.37e-231 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
Gaps: 14
US-10-051-909-32 (1-800) x US-09-938-842A-2254 (1-2205)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 1 ATGAAGGAGCGAGCTCTGTTGCTCTCGCCGCCCAATCGCAATTCTTACAGATGG 60
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleIysGluPheGlnLeuGlnAsn 95
Db 61 GACAATGCCACCATGCTGTGAGCTATGATTATCAACAAGACTTGAATCTA----- 114
QY 96 GluProThr--ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CCAACCTCTGTTCAAGTCTTGTGCTGTATGTCATTGATCGGTGCAACGCTCATC 171
QY 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134
Db 172 ACGACTTGCTCAGAGCAGCATATCTGATTGGCTCGGACAGACGCCCATGCTCATTTATCA 231
QY 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
Db 232 TCAGTTATGATTTCGTCTGCGGTTTGATTAATGTTGTGCTCCCAATGCTAATGTTCTG 291
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 292 TGCTTTGCTAGGCTTCTTAATGGGTTGTGTCGCGGCTCGCGGTTACACTGTGCTGTT 351
QY 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 352 TACATTTCTGAACCGCTCTCCGAGATCAGAGACAGTTAATACTCTCCCTCAGTTT 411
QY 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 412 CTGGCTCTGTGGAATGTTTGTTCATATCTGATGCTTTTCACTATGTCTCCCTGAGTGA 471
QY 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db 472 TCCCTAGCTGAGAGCATGCTCGGTGCTCTTCGATCCCTTCTCTTTATTTGTTT 531
QY 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIysGlyArgMetAla 254
Db 532 CTCACGGTGTATTATGTCGCCGAGTCTCCTCGTTGCTGTTAGTAAGAAGATGAGC 591


```

QY 308 AsplysglunlethrleuTyrglyProgluglunleuSerTrpIlealaArgPro 327
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 21 GAAAAAGATCAAAATTAAGTTGATGAGCAGAACAGGCGAGTCTGGGTGCTAGACCT 80
QY 328 SerlysglyProilemetleuGlySerValleuSerleualaSerArgHisGlySerMet 347
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 81 GTTGCTGACCAAAATCTGTGGC-----CTTGATCTAGGAAAGAACATG 128
QY 348 ValAsnGlnSerValProleuMetAspProIleValThrleuPheGlySerValHisGlu 367
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 129 GCAATCCAGC---AGCTAGTGAGCCCTCTAGTACCCCTTGAGTAGTACATGAG 185
QY 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrleuPheProAsnPheGlySerMet 387
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 186 AAGCTCCAGAAACAGG-----AGCACCTTTTCCACACTTGGGAGATATG 233
QY 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 234 TTCAGTGTGGGGGAAATCAGCCAGGAATGAAGATGGAGTGAAGAAAGCCTAGCCAGA 293
QY 408 AspAspGluGluTyrrAlaSerAspGlyAlaGlyGlyAspTyrglyAspAsnLeuHisSer 427
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 294 GAGGATGATGATGATGATCTGAT-----GCTGTGATCTCTGATGACAAATTCACAGAT 347
QY 428 ProleuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 348 CCATGATCTCAGCTCAACAAACAGAGCTGAT---AAGACATACCTCCCTCATGCCAT 404
QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrleu---GlyGluGlyLysAspGly 466
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 405 AGTAACCTTGCA---AGCATGAGGCAAGGATAGTCTTTTACATGGAATTCAGAGAACCC 461
QY 467 ValSerSerThrAspIleGlyGlyGlyTyrglnLeuAlaTrpLysTrpSerGlyLysGlu 486
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 462 ACTGCTAGTACTGGATGGTGGTGGTGGTGGCAGCTAGCATGAAATGCTGGAAGAGAG 521
QY 487 GlyGluAsnGlyArgGlyGluGlyGlyPheLysArgValTyrrLeuHisGlnGluVal 506
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 522 GGCCAGATGGAAGAAAGAGAGGTGGCTTCAAGAGAAATATATTACACCAAGATGGTGT 581
QY 507 ProGlySerArgArgGlySerIleValSerleuProGlyGlyLysAspValPheGluGly 526
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 582 TCTGATCTAGACGTGGTGTGTTGCTTCACTCCCT---GGCGGTGATTACCAACTGAC 638
QY 527 SerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 639 AGTGAGGTTGTACAGGCTGCTGCTGCTGAGTCAAGCTGCTCTTATATAGAGACCTT 698
QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 699 ATGCGTCAACGGCCAGTTGGACCAAGCTATGATTCATCCCTCTGAAAACAATTGCAAGCG 758
QY 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly 586
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 759 CCAAGTTGAGTGTCTTTTGAACCTGGGTGAAGCATGATGATGTGGGGGCGGA 818
QY 587 IleGlnIleleuGlnPheAlaGlyIleAsnGlyValleuTyrrThrProGlnIle 606
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 819 ATGCAATTTCTCAGCAGTCTCTGTATATAATGGGCTCTACTATACGCTCAATT 878
QY 607 LeuGluGlnAlaGlyValAlaValIleleuSerLysPheGlyLeuSerSerAlaSerAla 626
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 879 CTTGAGCAGGCAAGGTGGTTATCTTTCTTCAAGCCCTAGGCCCTTGCTTCTTCA 938
QY 627 SerIleleuIleSerSerleuThrThrleuLeuMetLeuProCysIleGlyPheAlaMet 646
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 939 TCCTTTCTTATTAGTGGGTGACAACTTGTGATGCTTCTTGATAGCCATTGCCATG 998
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheleuLeuGlyThrIleProIleleuIle 666
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 999 AGGCTCATGATATTTCAAGCAGAGACTTGTGCTCAGTACAAATCCCGTCTTAATA 1058
QY 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrleuAlaHisAla 686

```

```

Db 1059 GCAGCTCTTCATATTAAGTCCCTGGAGAGCTTGTGATTTGGGATCCACTGCAATGCA 1118
QY 687 leuLeuSerThrValSerValIleValTyrrPheCysCysPheValMetGlyPheGlyPro 706
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1119 TCAATCTCAACCATTAAGTGTATGTCATTTCTGTTCTTGTGATGGATTTGGACCA 1178
QY 707 IleProAsnIleleuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1179 ATTCCTAATATACTTTGTGACAGATCTTCCCACTCGAGTTCGTGCTCTGCAATTGCT 1238
QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrrSerleuProValMet 746
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1239 ATTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGACCTACACACTCCAGTTATG 1298
QY 747 leuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrrAlaValValCysLeuIleSer 766
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1299 CTCATTTCTGAGGCTCGCTGTGTTTGGATTTATGCTGTGCTGTGCTTCAATAGCA 1358
QY 767 PheValPheValPheleuLysValProGluThrLysGlyMetProleuGluValIleThr 786
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1359 TGGGTGTTGTCTTTTGAAGTTCAGAAACCAAGGCAATGCCACTGGAAGTATCAT 1418
QY 787 GluPhePheAlaValGlyAlaLysGln 795
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1419 GAGTCTTCTCTGTGGAACAAACAG 1445

```

RESULT 10

```

US-10-051-909-9
; Sequence 9, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B0163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-909-9

```

Alignment Scores:

```

Pred. No.: 2.27e-167 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 14 Gaps: 8

```

US-10-051-909-32 (1-800) x US-10-051-909-9 (1-1692)

```

QY 308 AsplysglunlethrleuTyrglyProgluglunleuSerTrpIlealaArgPro 327
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 21 GAAAAAGATCAAAATTAAGTTGATGAGCAGAACAGGCGAGTCTGGGTGCTAGACCT 80
QY 328 SerlysglyProilemetleuGlySerValleuSerleualaSerArgHisGlySerMet 347
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 81 GTTGCTGACCAAAATCTGTGGC-----CTTGATCTAGGAAAGAACATG 128
QY 348 ValAsnGlnSerValProleuMetAspProIleValThrleuPheGlySerValHisGlu 367
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 129 GCAATCCAGC---AGCTAGTGAGCCCTCTAGTACCCCTTGTGATGATCATGAG 185

```

QY 368 AsnMetProGlnAlaGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCAGAAACAGGA-----AGCACCTTTTCCACACTTTGGAGATATG 233
QY 388 PheSerValThrAspGlnHisAlaLysAsnGlyGlnTrpAspGlyGluAsnLeuHisArg 407
Db 234 TTCAGTGTGGGGAATCAGCCAGGAATGAGATGGGATGAGAAAGCCTAGCCAGA 293
QY 408 AspAspGlyGlyTyrAlaSerAspGlyAlaGlyLysPheGluAspAsnLeuHisSer 427
Db 294 GAGGGTATGATATATGCTCTGAT-----GCTGATGATCTGATGACATTTGCAGAGT 347
QY 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGlyLysAspIleValHisHisGlyHis 447
Db 348 CCATTTGATCTCAGCTCAACACACAGACTCTGAT--AAGACATACTCTCATGCCCCAT 404
QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGlyGlyGlyAspGly 466
Db 405 AGTAACCTTGCA--AGCATGAGGCAAGGTAGTCTTTTACATGGAATTCAGAGAACCC 461
QY 467 ValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGlyLysGlu 486
Db 462 ACTGTAAGTACTGGATGGTGGTGGTGGCAGCTAGCAATGGAATGCTGAAAGAGAG 521
QY 487 GlyGluAsnGlyArgLysGlyGlyGlyPheLysArgValTyrLeuHisGlnGlyVal 506
Db 522 GGGCCAGATGGAAGAAGAGAGGTGGCTTCAAGAGATATATTTACACCAAGATGGTGT 581
QY 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlyGly 526
Db 582 TCTGGATCTAGACGTGGGTGTGTGTTTCACTCCCT--GGCGGTGATTTACCACTGAC 638
QY 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
Db 639 AGTGAGGTGTACAGGCTGTCTGTCTGTGAGTCAAGCCTGCTTTATATGAGAGCTT 698
QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566
Db 699 ATGCGTCAACGGCCAGTTGGACCACTATGATTCATCCCTCTGAAACAATTGCAAAAGCG 758
QY 567 SerArgTrpLysAspLeuPheGlnProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db 759 CCAAGTTGAGTGTATCTTTTGAACCTGGGGTGAAGCAATGATGATGGGGGTGGA 818
QY 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIle 606
Db 819 ATGCAATTTCTCAGCAGTCTCTGTGTATAATGGGCTCTCTACTATACGCTCAATT 878
QY 607 LeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTTGAGCAGGCAAGGTGTGTGTATCTTTTCAAGCCTAGGCCCTGTCTACTTCTTCA 938
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTTCTTATTAGTGGGTGACCACTTGTGATGCTTCTGTATAGCCCATGGCATG 998
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCATGATATTTCAGGCAAGACTTGTCTGCTCAGTACAATCCCCGTCTAATA 1058
QY 667 AlaSerLeuValIleLeuValAlaSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATATTAGTCTCGGAAGTCTTGTGATTTGGGATCCACTGCAATGCA 1118
QY 687 LeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATTAAGTTATTTCTTATTTCTTTTGTTCATGGGATTTGACCA 1178
QY 707 IleProAsnIleLeuCysAlaGlyIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCCTAATATACTTTGTGAGAGATCTTCCACACCGAGTTGCTGTCTGCAATGCT 1238
QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746

Db 1239 ATTTGTGCCCTTACTTTTGGATCTGTGATATCATTTGTACCTACACACTCCAGTATAG 1298
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
Db 1299 CTCATTTCTGTAGGCCCTCGCTGTGTGTTTGTGATTTATGCTGTCTGTCTTCATAGCA 1358
QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTGTCTTTTGAAGTTCAGAAACCAAGGGCATGCCACTGGAATGATCATT 1418
QY 787 GluPhePheAlaValGlyAlaLysGln 795
Db 1419 GAGTTCTTCTGTGTGAGCAAAACAG 1445
RESULT 11
US-10-051-902-13
; Sequence 13, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR APPLICATION NUMBER: 2002-01-17
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-13
Alignment Scores:
Pred. No.: 3,36e-141 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
Gaps: 2
US-10-051-909-32 (1-800) x US-10-051-902-13 (1-1487)
QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481
Db 12 GAGGGTGGGAGGCGACTCAGCAGCACTGTATTTGGTGGGGGTGCGCACTCGCATGAAA 71
QY 482 TrpSerGluLysGlyGlyGlyLysAsnGlyArgLysGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGAGCGACAGGCGAGGATGGCAAGAGAGGCTTCAAAAGAACTACTTG 131
QY 502 HisGlnGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGTGGCCGACTCAAGAAAGGGCTGTGTTTCACTTCTGTGGGGGT 191
QY 522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGGCAGTGGGTTTATACATGCTGCTGTGTAAGCCACTCGGCT 251
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet--SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGGATCTTATGGAAGAGCGTATGGCGCGGCTCCAGCCATGATTCATCCA 311
QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlnProGlyValArgArg 579
Db 312 TTGAGGCGAGCTCCCAAAAGTTCAATCTGGAAGAATCTGTTGAACCTGTGTGAGGCGT 371
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyVal 599


```
US-10-051-902-15
; Sequence 15, Application US/10051902
; Publication No: US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIORITY FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIORITY FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-15

Alignment Scores:
Pred. No.: 3.14e-90 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
DB: 14 Gaps: 0

US-10-051-909-32 (1-800) x US-10-051-902-15 (1-1009)
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db 2 GAACCTGAGTGAAGCATGCACTGTTGCTGGCATAGATTAACAGATCCTGCAGAGTTT 61
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
Db 62 GCGGGTATCAATGAGTCTCTACTACACACCTCAGATCTTGAGCAAGCAGGTGCGG 121
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeu 633
Db 122 GTTCTTCTATCAACATGAGTGAAGTCTTCTCAGCATCTATTCTATTAGTGCCTTG 181
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 182 ACAACCTGTGATGCTGCCAGCATTTGGCATCGCCATGAGACTCATGATATGTCAGGA 241
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 242 AGAAGTTTCTCTCTTCAACAATCCCTGCTTATAGTAGCGCTAGCTGTGCTT 301
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 302 TTAGTGATGTTCTGATGTCGAAACCATGTCGACGTCGCTCTCAACGATCAGCGTC 361
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 362 ATCGTCTATTCTGCTTCTGTCATAGGGGTTGGGCTTATCCCAATATTCTTGGCG 421
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 422 GAGATTTCCTCCACCTCTGTCGTCGATCTGCATAGCCATCTGCGGCTAACCTTCTGG 481
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 482 ATCGGCGACATCATCGTACATACACTCTCCCGTATGCTCAATGCCATTGTTCTGCT 541
QY 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLys 773
Db 542 GAGGCTTCGCGCATATATGTCATCGTTGTGACTAGCTTTGTATTGCTTACATGAAG 601
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
```

```
|||||
Db 602 GTCCCTGAGACAAAGGCATGCCCTGAGGTTCATCACCGAGTCTTCTCTGCGGCA 661
QY 794 LysGln 795
Db 662 AAGCAG 667

RESULT 14
US-10-051-909-15
; Sequence 15, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIORITY FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIORITY FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-909-15

Alignment Scores:
Pred. No.: 3.14e-90 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
DB: 14 Gaps: 0

US-10-051-909-32 (1-800) x US-10-051-909-15 (1-1009)
```

```
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db 2 GAACCTGAGTGAAGCATGCACTGTTGCTGGCATAGATTAACAGATCCTGCAGAGTTT 61
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
Db 62 GCGGGTATCAATGAGTCTCTACTACACACTCAGATCTTGAGCAAGCAGGTGCGG 121
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeu 633
Db 122 GTTCTTCTATCAACATGAGTGAAGTCTTCTCAGCATCTATTCTATTAGTGCCTTG 181
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 182 ACAACCTGTGATGCTGCCAGCATTTGGCATCGCCATGAGACTCATGATATGTCAGGA 241
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 242 AGAAGTTTCTCTCTTCAACAATCCCTGCTTATAGTAGCGCTAGCTGTGCTT 301
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 302 TTAGTGATGTTCTGATGTCGAAACCATGTCGACGTCGCTCTCAACGATCAGCGTC 361
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 362 ATCGTCTATTCTGCTTCTGTCATAGGGGTTGGGCTTATCCCAATATTCTTGGCGG 421
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 422 GAGATTTCCTCCACCTCTGTCGTCGATCTGCATAGCCATCTGCGGCTAACCTTCTGG 481
```

QY 734 IIEGLYASPIIEIIEValThrTyrSerLeuProValMetLeuAsnAlaIIEGLYLeuAla 753
Db 482 ATCGCGACATCATCTGACATACACTCTCCCGTATGCTCAATGCCATTTGCTCGCT 541
QY 754 GLYValPheSerIIEtyrAlaValCysLeuIleSerPheValPheValPheLeuLys 773
Db 542 GGAGTCTTCGCAATATGCAATCGTTGTGTACTAGCCTTGTATTCGTTCATGAAG 601
QY 774 ValProGluThrIleGlyMetProLeuGluValIIEThrGluPhePheAlaValAla 793
Db 602 GTCCCTGAGACAAGGCGATGCCCTGAGGTCATCACCGAGTTCTTCTGTGCGGGCA 661
QY 794 LysGln 795
Db 662 AAGCAG 667

RESULT 15
US-10-369-493-47011
; Sequence 47011, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 47011
; LENGTH: 1374
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-47011

Alignment Scores:
Pred. No.: 3.97e-52 Length: 1374
Score: 587.50 Matches: 168
Percent Similarity: 36.64% Conservative: 102
Best Local Similarity: 22.80% Mismatches: 162
Query Match: 14.42% Indels: 305
DB: 12 Gaps: 12

US-10-051-909-32 (1-800) x US-10-369-493-47011 (1-1374)

QY 55 LysMetSerGlyAlaValLeuValAlaIIEValAlaSerIIEGLYAsnLeuLeuGlnGly 74
Db 4 AAAAAAGCAGTCMAATATATGCTTTATTTTCGAGCTCTTGAGCGCGCTTATATGCG 63
QY 75 TrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGln 94
Db 64 TATGATACCGAGATGATTCGCGAGCTATTATTATGAAAAAGAGTTAGGCTTA--- 120
QY 95 AsnGluProThrValGluGlyLeuIIEValSerMetSerLeuIIEGLYAlaThrIleVal 114
Db 121 ---AACGCGTTTACAGAGGCTTGTGTGTACGCTCCTGCTGTTGGGCGATATGGGC 177
QY 115 ThrThrPheSerGlyProLeuSerAspSerIIEGLYArgArgPrometLeuIleLeuSer 134
Db 178 TCAGAGAGCGCGCGCAAGCTGACTGACCGCTTCGAGAGAAAAAAGCAATTATGCGACCC 237
QY 135 SerIleLeuTyrPhePheSerGlyLeuIIEMetLeuTyrSerProAsnValTyrValLeu 154
Db 238 GCGCTGCTGTTTGTATAGCGGCTTGTGTGTGACACTGGCCCAATATACAGAGTCATG 297
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIIEGLYLeuAlaValThrLeuValProLeu 174
Db 298 GTGCTGTTTTCGATCATTTTGGAGCTTGACGTCGGAACATCGACGACATCGTACCCCTT 357

QY 175 TyrIleSerGluIIEAlaProSerGluIIEArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 358 TATTATCTGAAGTGGCGCCAAACATATAACCGGGCGCTGTCTCATCTGATCAGCTG 417
QY 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 418 ATGATCACGGTCGCGATCTCTTCTTCTTAC-----ATGTCAATTACATATTGCCGAT 471
QY 215 SerProAspTrpArgIIEMetLeuGlyValLeuAlaIIEProSerLeuPhePheGly 234
Db 472 GCCGAGCGTGGCGCTGATGCTTGATTTGCTGTGCCGTCAATGCTTCTGCTT--- 528
QY 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAla 254
Db 529 ATTGCATTTTGTATTATGCCGAGAGCGCGCTGCTGTTCACGAATGGCGAAGAAGC 588
QY 255 GluAlaLysIIEValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSer 274
Db 589 AAAGCGAAGAAATTTCTTGAAAAATTGCGTGCAACAAGATATT----- 633
QY 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIIEGluGluTyrIIEIIEGly 294
Db 633 ----- 633
QY 295 ProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIIEThrLeu 314
Db 633 ----- 633
QY 315 TyrGlyProGluGluGlyGlnSerTrpIIEAlaArgProSerLysGlyProIIEMetLeu 334
Db 633 ----- 633
QY 335 GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu 354
Db 633 ----- 633
QY 355 MetAspProIIEValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
Db 633 ----- 633
QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
Db 633 ----- 633
QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluTyrAlaSer 414
Db 634 -----GATCAGAA----- 642
QY 415 AspGlyAlaGlyLysAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db 642 ----- 642
QY 435 ThrGlyAlaGluGlyLysAspIIEValHisGlyHisArgGlySerAlaLeuSerMet 454
Db 643 -----ATACAT----- 648
QY 455 ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIIEGlyGly 474
Db 649 -----GATATA----- 654
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGlyLysGluGlyLysAsnGlyArgLysGluGly 494
Db 655 -----AAAGAAGCGGAAGCAAGATGAAGCGGTCTGAAGAG--- 693
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgLysSerIIE 514
Db 693 ----- 693
QY 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534
Db 693 ----- 693

```
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
Db 693 -----
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574
Db 694 -----CTGTTGAT 702
QY 575 ProGlyValArgArgAlaLeuValGlyValGlyLeuGlnIleLeuGlnIlePheAla 594
Db 703 CCATGGGTGCGCCAGCGCTTATTCAGAGTTGGACCTGCTTTTTCAGACAATTATC 762
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAlaVal 614
Db 763 GGAACGAATACGATCATCTACTATGCCCAAGACCCTTACAAACGTCGA----- 813
QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeu-----Ile 630
Db 814 -----TTCGA-----AACTCCGCTTCGATTTTAGGCACGCTCGAATC 852
QY 631 SerSerLeuThrThrLeuMetLeuProCysIleGlyPheAlaMetLeuMetAsp 650
Db 853 GGCAAGTCAATGTTCTCATGACATTA-----GTAGCGATTAAATCATCGAC 900
QY 651 LeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670
Db 901 AAGATTGGAAGAAGCCGTACTGCTATTCGGAATGCGGCATGTGATCAGCTTGATC 960
QY 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrIleValHisAlaLeuLeuSerThr 690
Db 961 GTTCTCGCTTAGTAATCTCTTTTTCATAACACTCCGCTCCCTCATGAGCAGCCGTC 1020
QY 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
Db 1021 ATTGTTTAGCGCGTGTATGCTGTCTTTGCGGTACGCTGGGACCGGTTGTGGGTG 1080
QY 711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730
Db 1081 ATGCTTCCTGAATTGTTCCCGCTTCACGTCAGAGGAATCGGACCGGTGTTGACCTTA 1140
QY 731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
Db 1141 ATGTTGACGTTGGACACTGATGTTTCATTAACTATCAATATTAAATGAAGCATC 1200
QY 751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal 770
Db 1201 GGAATCAGTATTATTCCTGATTATGCGCGCATCGGTATCATGCGTCTTATTGTC 1260
QY 771 PheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGlu 787
Db 1261 CGATTTTAAAGTGACAGACAAAGGAAGACCTTGAAGAAATTGAGCAG 1311
```

Search completed: January 6, 2004, 00:41:21
Job time : 748 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 19:06:39 ; Search time 122 Seconds
(without alignments)
2894.316 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPLDRR.....PLEVITFEFAVGAKQAAXA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool/US10051909/runat_05012004_094743_22329/app_query.fasta_1.967
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10051909 @CGN 1_1_56 @runat_05012004_094743_22329 -NCPU=6 -ICPU=3
-NO MAP -LARGEJOURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Issued_Patents_NA:*
- 2: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/BACKFILE1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3517	86.3	2824	4	US-09-291-922-1
2	2674	65.6	2601	4	US-09-291-922-7
3	1670.5	41.0	1692	4	US-09-291-922-9
4	1424	34.9	1487	4	US-09-291-922-13
5	943	23.1	1009	4	US-09-291-922-15
6	562	13.8	870	4	US-09-291-922-5
7	496	12.2	1853	4	US-09-291-922-23
8	494.5	12.1	2017	4	US-09-291-922-21
9	491.5	12.1	1914	4	US-09-291-922-19
10	489	12.0	1872	4	US-09-291-922-27
11	463.5	11.4	2089	4	US-09-291-922-25
12	412	10.1	441529	3	US-09-103-840A-1

13	368	9.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
14	355	8.7	2856	4	US-09-643-597-135	Sequence 135, App
15	355	8.7	2856	4	US-09-480-884A-135	Sequence 135, App
16	355	8.7	2856	4	US-09-542-615A-135	Sequence 135, App
17	355	8.7	2856	4	US-09-606-421B-135	Sequence 135, App
18	354.5	8.7	3000	2	US-08-928-692-9	Sequence 9, Appli
19	354.5	8.7	3000	4	US-09-339-972-9	Sequence 9, Appli
20	332.5	8.2	2592	4	US-09-591-025-8	Sequence 8, Appli
21	275.5	6.8	972	4	US-09-134-001C-1798	Sequence 1798, Ap
22	266.5	6.5	443	4	US-09-291-922-3	Sequence 3, Appli
23	265.5	6.5	1428	4	US-09-328-352-1580	Sequence 1580, Ap
24	256	6.3	918	4	US-09-107-532A-561	Sequence 561, App
25	247	6.1	2343	3	US-09-031-392-1	Sequence 1, Appli
26	247	6.1	2343	2	US-09-299-549-1	Sequence 1, Appli
27	247	6.1	2343	4	US-09-610-417-1	Sequence 1, Appli
28	235	5.8	155	4	US-09-313-294A-4597	Sequence 4597, Ap
29	234.5	5.8	510	4	US-09-291-922-11	Sequence 11, Appli
30	227.5	5.6	1467	4	US-09-328-352-2465	Sequence 2465, Ap
31	225.5	5.5	1521	4	US-09-252-991A-989	Sequence 989, App
32	218.5	5.4	1002	4	US-09-252-991A-1099	Sequence 1099, Ap
33	217.5	5.3	1371	4	US-09-328-352-1320	Sequence 1320, Ap
34	207.5	5.1	1230	4	US-09-252-991A-947	Sequence 947, App
35	205.5	5.0	603	4	US-09-252-991A-1063	Sequence 1063, App
36	205	5.0	2842	4	US-09-620-312D-574	Sequence 574, App
37	198.5	4.9	1344	4	US-09-252-991A-3992	Sequence 3992, Ap
38	198.5	4.9	1479	4	US-09-252-991A-3960	Sequence 3960, Ap
39	196.5	4.8	2102	2	US-08-647-397-1	Sequence 1, Appli
40	192	4.7	1896	3	US-08-501-572-6	Sequence 6, Appli
41	192	4.7	1896	3	US-09-040-444-6	Sequence 6, Appli
42	190.5	4.7	2871	4	US-09-252-991A-11431	Sequence 11431, A
43	190	4.7	1320	4	US-09-252-991A-10535	Sequence 10535, A
44	190	4.7	1473	4	US-09-252-991A-10639	Sequence 10639, A
45	187.5	4.6	1251	4	US-09-328-352-2549	Sequence 2549, Ap

ALIGNMENTS

RESULT 1
US-09-291-922-1
Sequence 1, Application US/09291922
Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
EARLIER FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 1
LENGTH: 2824
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (29)
FEATURE:
NAME/KEY: unsure
LOCATION: (622)
FEATURE:
NAME/KEY: unsure
LOCATION: (636)
FEATURE:
NAME/KEY: unsure
LOCATION: (638)
FEATURE:
NAME/KEY: unsure
LOCATION: (669)

```

? FEATURE:
? NAME/KEY: unsure
? LOCATION: (771)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (822)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (856)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (889)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (896)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (944)
?
US-09-291-922-1

```

Alignment Scores:	
Pred. No.:	0
Score:	3517.00
Percent Similarity:	95.18%
Best Local Similarity:	91.57%
Query Match:	86.31%
DB:	4
Length:	2824
Matches:	684
Conservative:	27
Mismatches:	34
Indels:	2
Gaps:	2

US-10-051-909-32 (1-800) x US-09-291-922-1 (1-2824)

QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	238	ATGGGGGGCGCCGTGATGCTGCCATCGCGGCTTCTATCGCAACTTGCTGCAGGCTGG	297
QY	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn	95
Db	298	GAGAAATGCGACAATTGCTGAGCCGCTCTGTACATAAAGAAGAAATTC AACCTGCAGAGC	357
QY	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	358	GAGCCTCTGATCGAAGGCCTCATCGCCCATGTTCTCATTTGGGGCAACAGTCATCACA	417
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSer	135
Db	418	ACATCTCCGGGGCCAAAGGCGTGACTGCGGTTGGTAGAGGCCCATGCTGCTCGCTCGCT	477
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	155
Db	478	GTCTCTACTTCGTGAGTGCGGTGATGCTTTGGCGCCCAATTGTATACATCTTGCTC	537
QY	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	538	CTCGCAAGGCTCATTTGATGGGTTCCGATCGGTTTGCGGTCACACTTGTCTCTCTAC	597
QY	176	IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe	194
Db	598	ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNTNGAACACGTTGCCGAGTTT	657
QY	195	SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer	213
Db	658	ATTGGGGTCAGNGAGGAGATGTTCTCTCTACTGCATGGTGTGGAGTCCCTCATG	717
QY	214	ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe	233
Db	718	CCCAAACTGATTGGAGGCTCATGCTTGAGTCTGTGCATCCCGTCACTTAAATTACTTT	777
QY	234	GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet	253
Db	778	GGACTGACTGCTTCTACTTGCTGATTCACCAAGGTGGCTGTGTAAGCAAGAGAGATG	837
QY	254	AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeu	273
Db	838	GCGGAGGCGAAGAGAGTGTNTGCAAAAGCTTCGGGGAGAGAGAGATGTCTCANGGAGANG	897

QY	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle	293
		::: :::	
Db	898	GCTCTTCTAGTTGAAGGTTGGGGGTCGGTAAAGATACACCTATTTNAGAGTACATCATT	957
QY	294	GlyProAlaThrGluAlaAspAspLeuValThrAspGlyAspIlyGluGlnIleThr	313
Db	958	GGACCTGCACCGACCGACCGCATGATCTTGTAACGTACGCGTGATAAGAACAAATCACA	1017
QY	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerIlyGlyProIleMet	333
Db	1018	CTTTATGGGCGCTGAAGAAGCCAGTCATGAGATTGCTCGACCTTCTAAGGACCCATCATG	1077
QY	334	LeuGlySerValLeuSerIleuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	1078	CTTGAAGTGTGCTTTCTCTTGCACTCTCGTCATGGAGCATGTGTGAACCAAGAGTGTACCC	1137
QY	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	1138	CTTATGATCCGATGTGACACTTTTGGTAGTGTCCATGAGAAATATGCTTCAAGCTGGA	1197

QY	374	GIYSerMetArgSerThrIleuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1198	GGAAGTATGAGGAGCACATGTGTTCCAAACTTTGGAAGTATGTTCAGTGCACAGATCAG	1257
QY	394	HISAlaIysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluIleTyrAla	413
Db	1258	CATGCCAAAAATGAGCAGTGGGATGAAGAATCTTCATAGGATGACGAGAGTACGCA	1317
QY	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1318	TCTGATGTCGACGAGAGTCACATATGAGGACAATCTCCATAGCCCATTTGCTGCCAGGCAG	1377
QY	434	AlaThrGlyAlaGluGlyLysAspIleValHisIleGlyHisArgGlySerAlaLeuSer	453
Db	1378	GCAACAGGTGCGGAAGGGAAGAACATTGTGCACCATGTCACCGTGGAAGTCTTTGAGC	1437
QY	454	MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1438	ATGAGAAGCCAAAGCCTTTAGGGGAGGTGAGATGTTGTAGCAGCACTGATATCGGT	1497
QY	474	GlyGlyTyrGlnLeuAlaTyrPlySerSerGluLysGlyGlyGluAsnGlyArgLysGlu	493
Db	1498	GCGGATGCGAGCTTGCTTGAATGTCAGAGAAGGAAGTGAAGATGTAGAAAGGA	1557
QY	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1558	GGTGGTTTCAAAAGAGTCTACTTGACCAAGAGGAGTTCCTGGCTCAAGAAAGGGCTCA	1617
QY	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTTCACTTCCCGGTGGTGGCGATGTTCTTGAAGGTAGTAGTTGTACATGCTGCT	1677
QY	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCTTTAGTAGTCAGTCAGCACTTTCTCAAAGGGTCTTGCTGAACCAAGCATGTCAGAT	1737
QY	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrPlyAspLeuPhe	573
Db	1738	GCTGCCATGTTCAACCATCTGAGGTAGCTGCCAAAGGTTCACGTTGAAAGATTGTTT	1797
QY	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe	593
Db	1798	GAACCTGAGTAGGCGGTGCCCTGTAGTCGGTGTGGAATTCAAGATCCCTCAACAGTTT	1857
QY	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGlyValAla	613
Db	1858	GCTGAATTAACGGTGTCTGTACTATATACCCCAAAATCTTGAGCAAGCTGTGTGGCA	1917
QY	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633
Db	1918	GTTATTTCTTCCAAATTTGGTCTCAGCTCGGCAATCAGATCCATCTTGATCAGTTCTCTC	1977

QY	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1978	ACTACCTTAATGCTTCCTTGCATTTGGCTTTGCCATGCTCTTAGATCTTTCGGA	2037
QY	654	ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	2038	AGAAAGTTTGTGCTGCTAGGACACAATTCCAATCTTGATAGCATCTCTAGTATCCTGGTT	2097
QY	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	2098	GTGTCCAATCTAATTGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT	2157
QY	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2158	ATCGTCTACTTCTGCTGCTTCGTTAGGGATTGGTCCCATCCCAACATTTATGTGCA	2217
QY	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp	733
Db	2218	GAGATCTTTCCAACCAAGGGTTGTTGGCCTCTGATTTGCCATTGTGCCATTACATTCGG	2277
QY	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2278	ATCGGAGATATCATCGTCACCTACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCG	2337
QY	754	GlyValPheSerIleTyrAlaValAlaValCysLeuIleSerPheValPheValPheLeuYs	773
Db	2338	GGTGTTCAGCATATATGCAAGTCGTATGCTTGAATTCCTTGTGTTCCTTCTTAAG	2397
QY	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793
Db	2398	GTCCCTGAGACAAGGGGATGCCCTTGAGGTTATTACGAATCTTTGCAGTTGTGGCG	2457
QY	794	LysGlnAlaIleAlaIleAlaYsAla 800	
Db	2458	AAGCAAGCGGCTGCAAAAGCC 2478	

```

RESULT 2
US-09-291-922-7
; Sequence 7, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-7

```

Alignment Scores:	
Pred. No. :	2.46e-264
Score:	2674.00
Percent Similarity:	82.20%
Best Local Similarity:	69.88%
Query Match:	65.62%
DB:	4
	Gaps: . 8
	Length: 2601
	Matches: 522
	Conservative: 92
	Mismatches: 115
	Indels: 18

US-10-051-909-32 (1-800) x US-09-291-922-7 (1-2601)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ATGAAGTGCCCGTCTGTGCTATTGCCGCTTCATTTGGTAATTCCTCCAAGATGG 234

QY	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnAsn	95
Db	235	GATAAATGCTACCATCCCGGGCTAATGTTACATTAAAGAAAGACCTTGCTTTGGGAACA	294
QY	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	295	-----ACTATGGAAGGCTTGTTGGTGGCATGTCCTGATTTGAGCAACGGTATACACC	348
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSerSer	135
Db	349	ACATGCTCTGCTCCTATAGCCGATTGGCTCGGTCCGCAACCATGATGATATCTCATCT	408
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPseProAsnValTyrValLeuLeu	155
Db	409	GTGCTCTATTCTTGCGTGGTTTGGTGATGCTGTGTCCTCCCAATGTGATGTGTGTC	468
QY	156	LeuAlaArgPheValAspGlyPheGlyTyrIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	469	TTGGCGAGGCTACTTATGATGGATTGGATGGCTTGCTGTGACTCTTGTCCTCCGGTCTAT	528
QY	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db	529	ATATCTGAACCGCCCGCTCTGAATAAGGGGGTGTGAATACGCTTCTCTACGTTCCAGT	588
QY	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Db	589	GGCTCTGAGGAATGTTTTTGTGCTACTGATGTTTTTGGCATGTCAATGATGCCCGCG	648
QY	216	ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db	649	CCTAGCTGAGGCTCATGCTTGGGGTCTGTCTATCTCTCTCTGTATTTGCATTG	708
QY	236	ThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMetAlaGlu	255
Db	709	ACCATTTTTCTTGCCCGAGTCTCTCGGTGGCTGTCAACAAGGAGATGCTCGAG	768
QY	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
Db	769	GCTAAGAAGGTGCTCCAAGATTGCGCGAAGGAGGATGTGCAGGCGAGATGGCATTG	828
QY	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
Db	829	CTGCTGAAGGTCTCGGATTTGGGGGTGATACATCTATCGAAGAGTACATATTTGCCCT	888
QY	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
Db	889	GCTGACGATGTGGCTGATGTCATGAACATGCACAGAAAGATAAATTTCATTATAT	948
QY	316	GlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMetLeuGly	335
Db	949	GGATCCCAAGCAGGCTTTCTTGTTATCAAACTGTCACTGCACAGAGTTCTATTGGC	1008
QY	336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db	1009	-----CTTGGCTCACACCATGGAAGCATCATCAACCAAGAGATGCCCTCATG	1056
QY	356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Db	1057	GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAAGCTCCCGAGACAGGACAAGA	1116
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1117	GGAAGCATGCGAAGCACTCTGTTCCAAATTTCGAAGCATGTTTCAGCACACTGCTGAGCCG	1176
QY	394	HisAlaLysAsnGluGlnTyrAspGlyGluAsnLeuHisArgAspAspGlyGluTyrAla	413
Db	1177	CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAGAAGGAGCTGAGGACTACATG	1236
QY	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1237	TCAGATGCACACCCGTGGGACTCCGATGATTAATTGGACAGTCTTAAATCTCACGCCAA	1296
QY	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--	452

434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452
 Db 1237 TCAGATGCAACCCGCGGGGACTCCGATGATTAATTGCACAGTCCCTTTAATCTCAGCCAA 1296
 QY

Db 1297 ACAACAAGCCTTGAA---AAAGACTTAACCTCCTCCTCCCTTCCCATGGCAGTATCCTTGGC 1353
QY 453 SerMetArgArgGlnThrLeuLeuGlyGlyGly--GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGCGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
QY 472 IleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGlnLysGlyGlyGlyAsnGlyArg 491
Db 1414 ATTGGTGGTGGCTGGCAACTGGCAATGGAAATGGACTGATAAA---GGTAGAGATGAAAA 1470
QY 492 LysGlnGlyGlyPheLysArgValTyrLeuHisGlnGlnGlyValProGlySerArgArg 511
Db 1471 CAACAAGAGAGGTTTAAAGGATTATTTACATGAGAGGAGGAGTTCTGCATCTCGTCGT 1530
QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGlnGlySerGlnPheValHis 531
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTGTCCAG 1572
QY 532 AlaAlaIleLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db 1573 GCTGCTGCCTTGTTAAGCCCAACCGCTCTTACTCCAAAGAGCTTATTGATGACACCCA 1632
QY 552 SerAspAlaAlaMetValHisProSerGlnValAlaAlaLysGlySerArgTrpLysAsp 571
Db 1633 GTTGGGCGCTGCAATGGTTCAACCACTGAGACAGAGCTTCAAGGGGCCCAAGTTGGAAGCT 1692
QY 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTTCTTGAACCAAGGGGTTAAGCATGCAATTGGTTGTGAGATTGGAATACAAATACTCTCAG 1752
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGly 611
Db 1753 CAGTTTTCAGGATAAATGGGGTTCTATATTATACACACCTCAATCCTTGAAGAGCGCGGT 1812
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTTCTTTCAGATATATAGGCATTGGCTCAAGATCGGCATCATTCCTTATCAGT 1872
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGATGTT 1932
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAAGAGGAGGAGTGTACTACTACTAACAATCCCGTGTGATTTGTCTACTCATTAATT 1992
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGGTTCATTGGAAGCGCTGTAAATTTGGCAATGTCGCCCATGCAACAATCTCAACAGTA 2052
QY 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TGGGTGTGGTTTATTTCTGCTGCTTGTGATGGGTATGAGCAATTCCAACAATCCTT 2112
QY 712 CysAlaGlnIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTTCCCACTAGGGTGGCTGCGCTCTGCATTGCTATCTGTGCATTAGTG 2172
QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCGTGATTGGAGACATCATCATCACTACTGCTGCTGATGCTCGGCTCTTGA 2232
QY 752 LeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPhe 771
Db 2233 CTTGGTGTGTATTGCGCATTTACGACAGTTGTTGTTTCACTCTCGTGATATTGTGTTT 2292
QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db 2293 TTGAAGGTTCCAGAAACAAGGGAGTCCCGCTTGAAGTCACTCTGAATTCTTTCTGTT 2352
QY 792 GlyAlaLysGlnAlaAlaAla 798

Db 2353 GGAGCAAGCAGGCTGCTTCT 2373
RESULT 3
US-09-291-922-9
; Sequence 9, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
; US-09-291-922-9
Alignment Scores:
Pred. No.: 1.53e-161 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 4 Gaps: 8
US-10-051-909-32 (1-800) x US-09-291-922-9 (1-1692)
QY 308 AsplysGlnGlnIleThrLeuTyrGlyProGlnGlnGlyGlnSerTrpIleAlaArgPro 327
Db 21 GAAAAAGATCAAAATTAAGTTGTATGAGCACCAAGCAAGCCAGTCTGGGTGTAAGACT 80
QY 328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db 81 GTTCTGAGACCAATTCTGTTGGC-----CTGTATCTAGAAAGAAAGCATG 128
QY 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGln 367
Db 129 GCAAAATCCAAGC--AGTCTAGTGACCCCTTAAGTGAACCCCTTGTGATGTAACATGAG 185
QY 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCCAAGAAACAGGA-----AGCACCTTTTCCACACATTTGGGAGTATG 233
QY 388 PheSerValThrAspGlnHisAlaLysAsnGlnGlnTrpAspGlnGluAsnLeuHisArg 407
Db 234 TTCAGTGTGGGGAATCAGCCAAGAAATGAAGATTGGGATGAGAAAGCTAAGCCAGA 293
QY 408 AspAspGlnGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlnAspAsnLeuHisSer 427
Db 294 GAGGCGATGATTAATGCTCTGAT-----GCTGCGATTTCTGATGACAAATTTGCCAGAGT 347
QY 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGlnGlyLysAspIleValHisHisGlyHis 447
Db 348 CCATTGATCTACGTCAAACAACAGAGTGGAT--AAGACATACCTCTCATGCCCAT 404
QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu--GlyGlnGlyGlyAspGly 466
Db 405 AGTAACCTTGCA--AGCATGAGGCAAGGATGCTTTTACATGAAATTGAGAGAACCC 461
QY 467 ValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGlnLysGlu 486
Db 462 ACTGCTAGTACTGGGATGCTGCTGCTGGTGGCAGCTAGCATGGAATGCTGGAAGAAGAG 521
QY 487 GlyGluAsnGlyArgLysGlnGlyGlyPheLysArgValTyrLeuHisGlnGlnGlyVal 506

```
Db 522 GGGCCAGATGGAAGAAGAGGTGGCTTCAAGAGATATATTACACCAAGATGGTGT 581
QY 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGluGly 526
Db 582 TCTGGATCTAGACGTGGGTCTGTGGTTTCACTCCCT--GGCGGTGATTACCACTGAC 638
QY 527 SerGluPheValHisAlaAlaIleuValSerGlnSerAlaIleuPheSerIleGlyLeu 546
Db 639 AGTGAGGTGTGACAGGCTGCTCTGTGTAGTCAAGCTGCCCTTATATAGAGACCTT 698
QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaGly 566
Db 699 ATGCGTCAACGCGCAGTTGGACCAAGCTATGATTCATCCCTGAAACAATTGCAAAAGG 758
QY 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaIleuValGlyValGly 586
Db 759 CCAAGTTGGAGTGATCTTTTGAACCTGGGGTGAAGCATGCAATTGATGTGGGGTGGGA 818
QY 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrrTrpProGlnIle 606
Db 819 ATGCAAAATCTTCAGACAGTCTCTGTATAAATGGGGTCTCTACTATACGCTCAAAAT 878
QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTTGACAGCAGCAGGTGTGGTATCTTCTTCAAGCCTAGCCTTGGTTCTACTTCTTCA 938
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTTCTTATTAGTGGCGGTGACAACCTTGTGATGCTTCTGTATAGCCATTGCCATG 998
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCATGATATTTCAGGCAAGAAGACTTGTGCTCTCAGTACAAATCCCCCTCTAATA 1058
QY 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATATTAGTCTCGGAAGTCTGTGATTTGGATTCACCTGCAATGCA 1118
QY 687 LeuLeuSerThrValSerValIleValTyrrPheCysCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATTAGTGTATGTCTATTCTTCTTGTGATGGATTGGACCA 1178
QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCCATATATACTTGTGACAGATCTTCCCACTCGAGTCTGTGCTCTGCAATGCT 1238
QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrrSerLeuProValMet 746
Db 1239 ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGACCTACACACTCCAGTTATG 1298
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrrAlaValValCysLeuIleSer 766
Db 1299 CTCAATCTGTAGGCTCGCTGGTGTGTTTGGTATTATGCTGTGCTTCAATGCA 1358
QY 767 PheValPheValPheLeuValValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTTGTCTTTTGAAGTTCAGAAACCAAGGCGCATGCCACTGGAAGTGAATCA 1418
QY 787 GluPhePheAlaValAlaAlaGlyGln 795
Db 1419 GAGTCTTCTCTGTGCGAGCAAAACAG 1445
```

RESULT 4

```
US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
```

```
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-13

Alignment Scores:
Pred. No.: 2.61e-136 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 4 Gaps: 2
```

US-10-051-909-32 (1-800) x US-09-291-922-13 (1-1487)

```
QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyIleTrpGlnLeuAlaTrpLys 481
Db 12 GAGGTGGGAGGAGCATCAGACAGCACTGATGTGGGGGTGCAACTGCATGGA 71
QY 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyPheLysArgValTyrrLeu 501
Db 72 TGGTCGAGCGACACAGCGCAGGATGGCAAGAAGAGAGGCTTCAAAAGATCTACTTG 131
QY 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGTGGCGGCACTCAAGAAGGGCTCTGTGTTTCACTTCTGTGGGGCT 191
QY 522 AspValPheGlu--GlySerGluPheValHisAlaAlaIleuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGCGAGTGGTTTATACATGCTGCTTGTGTAAGCCACTCGGCT 251
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGATCTTATGGAAGAGCTATGGCGCGGTCCAGCCATGATCATCCA 311
QY 560 SerGluValAlaAlaGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGGAGGAGAGCTCCCAAGGTTCAATCTGGAAGATCTGTTGAACCTGTGTAGGGCT 371
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTTGCTCGGTGGTGAATTCAGATGCTTCAGCAGTTTGTGGAATTAATGAGTT 431
QY 600 LeuTyrrTrpProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTATACCTCTCAAAATTCGAGCAAGCTGTGTGGCTGTTCTTCTTCAATCTT 491
QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCCTCAGTTCAAGCATCAGCATCCATCTTGATCAGTTCTCTCACCACTTACTATGCTC 551
QY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATTTGGTGTAGCCATGAGACTTATGATATATCTGGAAGAAGGTTCTGTACTG 611
QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAAATTCATCTTATGATAGCATCCCTAATTTTGGGTGTGTCATATGATCAAC 671
QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrrPheCys 699
Db 672 TTGAGTACGGTGGCCGCGCTGTCTCTCCACAGTTAGCGTCATTTGTCTACTTGTCTGC 731
QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTATGAGGCTTTGGCCGATCCCAACATTTATGTGACAGATTTTCCCAACAGGA 791
```

QY 720 VALARGGLYLEUCYSILEALALECYAALAPHETHRPHETRPILEGLYASPILEILEVAL 739
| | | | | : : : : :
DB 792 GTCGGTGTGTCTGCATCCGTAATTGGCCCTCACATCTTGATTTGTGACATTATTGTT 851
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
| | | | | : : : : :
DB 852 ACCTAAGCCTGCTGCTGATGCTGAATGCTATTGTTAGCGGGTGTCTTTGGTATATAT 911
QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
| | | | | : : : : :
DB 912 GCAGTGGTTGCTGCATTGCGCTTGTGTCTACTAAAGGTCCAGAGACAAAGGC 971
QY 780 MetProLeuGluValIleThrGluPhePheAlaValAlaLysGlnAlaAla 798
| | | | | : : : : :
DB 972 ATGCCCTCGAGTCATCACCGAGTCTTTGCGGTTGGGCGAAGCAAGCGCAGGCC 1028

RESULT 5

US-09-291-922-15
; Sequence 15, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: *Triticum aestivum*
US-09-291-922-15

Alignment Scores:
Pred. No.: 3.56e-87 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
DB: 4 Gaps: 0

US-10-051-909-32 (1-800) x US-09-291-922-15 (1-1009)

QY 574 GluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
| | | | | : : : : :
DB 2 GAACCTGAGTGAAAGCATGCACTGTTGTTGGCATAGATTACAGATCCTGCAGCAGTTT 61
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
| | | | | : : : : :
DB 62 GCGGGATCAATGAGTCTCTACTACACACCTCAGATATTGAGCAAGCAGGTGCGG 121
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
| | | | | : : : : :
DB 122 GTTCTTCTATCAAAACATTTGACTTAAGCTCTTCTCAAGCACTTATTTAGTGCCTG 181
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
| | | | | : : : : :
DB 182 ACAACCTTGCTGATGCTCCAGCATTTGGCATCGCCATGAGACTCATGATATGTGAGA 241
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
| | | | | : : : : :
DB 242 AGAAGTTTCTCTCTTCAACAATCCCTGCTGTGATAGTAGCGCTAGCTGTCTGTT 301
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
| | | | | : : : : :
DB 302 TTAGTAAATGTTCTGATGCGAACCATGTGACGCTGCGCTCTCAACGATCAGCGCTC 361

QY 694 ILeValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
| | | | | : : : : :
DB 362 ATCGTCTATTCTGCTTCTTCGTCATGGGGTTGGCCCTATCCAAATATTTCTGCGCG 421
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
| | | | | : : : : :
DB 422 GAGATTTTCCCACTCTGTGCGGTGCGATCTGACATGACCATCTGCGCGCTTCTG 481
QY 734 ILeGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
| | | | | : : : : :
DB 482 ATCGGCACATCATGACATACACTCTCCCGTGAATGCTCAATGCCATTGGTCTGCT 541
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
| | | | | : : : : :
DB 542 GAGTCTTCGCATATATGCCATCGTTGTGTACTAGCCTTGTATTCGTCTACATGAAG 601
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
| | | | | : : : : :
DB 602 GTCCCTGAGACAAAGGCATGCCCTGAGGTCAACCGAGTTCTTCTGCGGGCA 661
QY 794 LysGln 795
| | | | |
DB 662 AAGCAG 667

RESULT 6

US-09-291-922-5
; Sequence 5, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 870
; TYPE: DNA
; ORGANISM: *Oryza sativa*
US-09-291-922-5

Alignment Scores:
Pred. No.: 3.89e-48 Length: 870
Score: 562.00 Matches: 103
Percent Similarity: 93.75% Conservative: 17
Best Local Similarity: 80.47% Mismatches: 8
Query Match: 13.79% Indels: 0
DB: 4 Gaps: 0

US-10-051-909-32 (1-800) x US-09-291-922-5 (1-870)

QY 669 leuValIleLeuValIleValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeu 688
| | | | | : : : : :
DB 8 GTTCTAACCTTGAATCTGCTCAATATTTGATGTGGGACCATGTTCAATGCTCACTG 67
QY 689 SerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708
| | | | | : : : : :
DB 68 TCCACAGTCAGTCACTACTCTTCTTCTTGTGCATGGGTTGGGCTATTTCCA 127
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
| | | | | : : : : :
DB 128 AACATTCTGTGAGAGATTTCGCCGACCAACCGTTGTCATCTGCAATAGCCATCTGT 187
QY 729 AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748
| | | | | : : : : :
DB 188 GCCCTAACATTTCTGATCGGTGATATCATTTGACATACACCCCTCCCGTATGCTCAAC 247

[illegible]

RESULT 7

```

US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23

```

Alignment Scores:

Pred. No.:	8.71e-41	Length:	1853
Score:	496.00	Matches:	165
Percent Similarity:	36.39%	Conservative:	109
Best Local Similarity:	21.91%	Mismatches:	189
Query Match:	12.17%	Indels:	290
DB:	4	Gaps:	18

US-10-051-909-32 (1-800) x US-09-291-922-23 (1-1853)

Dy	59	AAlaValleuValAlaIleValAlaSerileGlyAsnLeuLeuGlnGlyTrpaspAsnaLa	78
		::: ::: ::: :::	
Dd	257	GCTTTTCCTTGCTATGCTGGCCCTCCATGCATTCCATCTTGTGTTATGATATTGA	316
		::: :::: ::: ::: :::: :::: ::	
Dy	79	ThrIleAlaAlaAlaValLeuTyrrileYslYsgluPheGlnLeuGlnAsnGlu-----	96
		::: :::: ::: ::: :::: :::: ::	
Dd	317	GTCATGACGTGGAGCAGCACCATATACATAAAAGGCACTGAAAGTCTCGAGCAGCAATC	376
		::: :::: ::: ::: :::: :::: ::	
Dy	97	ProThrValGluglyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThr	116
		::: ::: ::: ::: :::: :::	
Dd	377	GAGATCCTGCTCGGAATCATCATCAACTATACTCTTGATAGGCTCATGTCTC-----	427
		::: ::: ::: ::: :::: :::	
Dy	117	PheSerglyProLeuSeraspSerileGlyArgArgPrometLeuIleLeuSergSerile	136
		::: :::: :::: :::	
Dd	428	---GCCCGCAGAACCCTCCGACTGGATAGGTCCCCGTTACAACGATTGTTTTCCGCCGCAACC	484
		::: :::: :::: :::	
Dy	137	LeutyrrPhePheSerglyLeuIleMetLeutyrPserProAsnValTyrrValLeuLeu	156
		::: :::: :::: :::	
Dd	485	ATCTTCTTTGTCGGAGCACCTTCTCATGGGTTTCTCCCAATPATTTCTTTTCATGTTT	544
		::: :::: :::: :::	
Dy	157	AlaArgPheValaspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrrile	176
		:::: :::: :::	
Dd	545	GGCCGTTTCGTGCTGGCATTGGCATGGCTACGCCCTCATGATAGCCCCCGTCAACC	604
		:::: :::: :::	
Dy	177	SerGluIleAlaProSerglyLeuIleArgGlyLeuLeuAsnThrLeuProGlnPheSergly	196
		::: :::: :::: :::	
Dd	605	GCCGAGGTCTCCCCGACCTCTCTCTGCTGCTTCTCATCTTCTCCCTGAGTAATTATT	664
		::: :::: :::: :::	

QY		197	SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSer---LeuSerProSer	215
Db		665	AATGAGGCGATATTAAATGGATACATATCAACTATGATTTTCGAAGCTGACACTAAAG	724
QY		216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db		725	GTCGGATGGCGAATGATGCTTGAGTTGGTGCATAACCTCCGTA--CTCCTAACAGTA	781
QY		236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255
Db		782	GGAGTGTGGCGGATGCCGAGTCCCCAAGGTGGCTGTGATGAGGGGTCGTTGGCAGAG	841
QY		256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
Db		842	GCAAGAAAAAGTCTTAACAAA-----	862
QY		276	LeuLeuGluGluLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
Db		862	-----	862
QY		296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
Db		863	-----ACCTCAGACAGCAGAAGAGAG-----	883
QY		316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
Db		883	-----	883
QY		336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db		884	GCCCAACTAAGGCTAGCG-----	901
QY		356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySer	375
Db		902	-----GAATCAAACAAGCCGACGGATC	925
QY		376	MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla	395
Db		926	CCCAGAGTTGC-----	937
QY		396	LysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAsp	415
Db		938	-----AACGACGAC-----	946
QY		416	GlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThr	435
Db		946	-----	946
QY		436	GlyAlaGluGlyLysAspIleValHisIleGlyHisArgGlySerAlaLeuSerMetArg	455
Db		947	-----GTCGTTCAAGTAAT	961
QY		456	ArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGlyGly	475
Db		962	AAACCAAGC-----AACGCTGAAGCTGTA-----	985
QY		476	TyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgGlySerGluGly	495
Db		985	-----	985
QY		496	PheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleVal	515
Db		985	-----	985
QY		516	SerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeu	535
Db		985	-----	985
QY		536	ValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAla	555
Db		985	-----	985

985 ----- 985

```
QY 556 MetValHisProSerGluValAlaAlaIleGlySerArgTrpLysAspLeuPhe----- 573
Db 986 -----TGGAAAGAGCTCTTCTCTAT 1006
QY 574 ---GluProGlyValArgAlaLeuValGlyValGlyIleGlnIleLeuGln 592
Db 1007 CCAACGCCCGCAATTCGTACATCGTAATCGCTGCCCTGGTATTCATTCTCCAAACA 1066
QY 593 PheIleGlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGlnAlaGlyVal 612
Db 1067 GCGTCGGCGGTAGACGCCGTCTTGTGTACAGCCCGAGATCTCGAAAGCTGGGATT 1126
QY 613 AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSer 632
Db 1127 ACAAGACAGACGATAGCT-----CTTGCAACCGTGGCCGTTGGATTTC 1171
QY 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652
Db 1172 GTTAGACCGGTTCATCTTGGCG-----GCTACGTTTACGTTGGACCGCGTG 1219
QY 653 GlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 1220 GGTGTGCGTCGCTGTGTATGTAGTGTGCGCGCATGGTGTCTCTCTCAGCGTT 1279
QY 673 ValValSer---AsnLeuIleAspLeuGly-----ThrLeuAlaHisAlaLeuLeuSer 689
Db 1280 GCGATTCAGCCTCACTGTTATGTATCATTCGAGAGAAATGATGTGGCCGTTGATCG 1339
QY 690 ThrValSer---ValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708
Db 1340 AGCATAGCCATGCTGTGTGCTACGTGGCCACGTTCTCCATCGGTGCGGTCCTCATCG 1399
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
Db 1400 TGGGTCTATAGTTCTGAGATCTTCCGTTGAGGCTGCGGCGCARGGTGGCGCGCGGA 1459
QY 729 AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748
Db 1460 GTTCCGGTGAATAGGACCACCTAGCGCGGTGTCTCAATGACTTTTCTCTCCCTCATAAGA 1519
QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheVal 768
Db 1520 GCCATCACTATTGGTGAGCTTCTTCTTATTGTGCAATGCTACTGTTGGTGATTA 1579
QY 769 PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr---Glu 787
Db 1580 TTCTTTTACACCGCTCTTGCCCTGAGACCGGGGAAAAACGCTCGAAGACATGGAAGGCTCT 1639
QY 788 PhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 1640 TTTCGTACTTTTAGTCCAAATCCAACGCCAGCAGAGCT 1678

RESULT 8
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinsley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
```

```
US-09-291-922-21
Alignment Scores:
Pred. No.: 1.44e-40 Length: 2017
Score: 494.50 Matches: 185
Percent Similarity: 35.71% Conservative: 101
Best Local Similarity: 23.10% Mismatches: 206
Query Match: 12.13% Indels: 309
DB: 4 Gaps: 22

US-10-051-909-32 (1-800) x US-09-291-922-21 (1-2017)
QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCysSerSerGln 43
Db 30 TTACACTCGACCCGCTACTGTACACGGCCCA---GAGCGAGCCTCTCTCTCTGCA 86
QY 44 GluProValThrSerAspAspIleLeuGluAsp-----LysMetSerGly--- 58
Db 87 CCACCGAGATGGCTTCCCGCGCGCTGCGGAGCGCTCGCGCGAAGAGAGGGCAAC 146
QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 147 GTCCGCTTCGCTTCGCTTCGCCATCTCGCTTCATGACCTCCATCTCTCGGCTAC 206
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 207 GATATCGGGTGATGAGCGGGGCGCTCGCTGTACATCAAGAAGACTTCAACATC--AGT 263
QY 96 GluProThrValGlu-----GlyLeuIleValSerMetSerLeuIleGlyAlaThr 112
Db 264 GACGGGAAGGTGAGGTTCTCATGGGCATACGTGAACCTTACTCGCTCATCGCTCTTC 323
QY 113 IleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIle 132
Db 324 GCG-----GCGGGGCGAGCGTCCGATCGGATCGCGCGCGGTACACCATCGTG 371
QY 133 LeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyr 152
Db 372 TTGCGCGCGCTCATATTCTTCCGGGGGSGTCTCATGGGGTTGCGCGCTCACTACGCC 431
QY 153 ValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuVal 172
Db 432 ATGCTCATGTTCCGCCGCTTCGTGCGCGGCAATCGCGCTACGCGCTCATGATCGCG 491
QY 173 ProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuPro 192
Db 492 CCGGTGTACACCGCGAGAGGTGTGCGCGCGCTGCGCGCTTCTTCACTGCTTCCCG 551
QY 193 GlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeu 212
Db 552 GAGGTGTTCATCACTTCGGCATCTGCTCGGTAACGCTCGAATATGCTTCTTCCCG 611
QY 213 SerPro--SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231
Db 612 TTGCGGTGAACCTCGGGTGGCGCATCATGCTCGGCAATGCGCGCGCTTCCGTG-- 668
QY 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGly 251
Db 669 CTGCTCGCGCTCATGTGCTCGCATGCGCGAGTCGCGCGGTGCTGCTCATGAAGGA 728
QY 252 ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGly 271
Db 729 CGCCTCGCGAGCGCAAGGTGTGTGAGAAAGACC----- 764
QY 272 GluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyr 291
Db 765 -----TCCGACACG----- 773
QY 292 IleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGln 311
Db 774 -----GCGAGAGAGCGCGCGAGCGCTG----- 797
QY 312 IleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyPro 331
```

```
Db      797 -----797
QY      332 IleMetLeuGlySerValIleuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351
Db      797 -----797
QY      352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371
Db      797 -----797
QY      372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391
Db      797 -----797
QY      392 AspGlnHisAlaLysAsnGlnGlnTrpAspGluGluAsnLeuHisArgAspAspGluGlu 411
Db      797 -----797
QY      412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431
Db      797 -----797
QY      432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAla 451
Db      797 -----797
QY      452 LeuSerMetArgArgGlnThrLeuLeuGlyGluGlyLysAspGlyValSerSerThrAsp 471
Db      798 -----GCCGAC 803
QY      472 IleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyLysAsnGlyArg 491
Db      804 ATCAAGCCGCC-----815
QY      492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db      816 -----GCCGCATCCCTGAGGAGCTCGAC 839
QY      512 GlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGluGlySerGluPhe 529
Db      840 GGCGACGTGTGACCGCTCCCAAGAGAGGAGCGGAAC-----878
QY      530 ValHisAlaAlaLysValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPro 549
Db      879 -----GAGAAAG 884
QY      550 ArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrp 569
Db      885 CGGGTG-----TGG 893
QY      570 LysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db      894 AAGGAGCTCATCTGTCCCGGACCCCGGCATGCGGCGCATCTGTGTCCGGGATCGGC 953
QY      587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIle 606
Db      954 ATCCACTTCTTCCAGCATGCGGTGGGCATTCACTCCGTCGTTCTTACAGCCCTCTCGTG 1013
QY      607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db      1014 TTCAAGAGCCCGGATTAAAG-----AACGACAAACACTTCTTGGCACCACCTTGGCCG 1067
QY      627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db      1068 TTCCGTGTACCAAGAGCTTTTCATCTTGTG-----GGCACT 1106
QY      647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db      1107 TTCTTCATCGACGCGCTGGGCGGCGCGCTGTGCTGGGACAGCAGGCGGGATATC 1166
QY      667 AlaSerLeuVal-----IleLeuValValSerAsnLeuIleAspLeuGly 681
```

```
Db      1167 CTCTCCCTCATCGGCTCGGGCGCGGCTACACCGTCTCGGCCAGCACCCGCGCAAG 1226
QY      682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701
Db      1227 ATACCTTGGGCCATCGCGCTTAAGCATCGCCTCACCTCGCTACGCTTCCTCTCTCC 1286
QY      702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg 721
Db      1287 ATCGGCTTGGCCCATCACGTGGGTGTACAGCTCGAGATCTTCCCGCTCAGGTGCGC 1346
QY      722 -----GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db      1347 GCGCTGGGCTGCTCGCTCGGCGCTCGCCCAACCGCTCACAGCGCGCTCATCTCCATG 1406
QY      740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db      1407 ACCTTCCTG-----TCGCTGTCCAAGGCCATCACCATCGGCGGACGCTTCTCTCTAC 1460
QY      760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db      1461 TCCGCGATCGCGCGCTCGCTGGGTGTCTTCTTACACCTACCTCCCGGAGACCCGCGGC 1520
QY      780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
Db      1521 CGGACGCTGAGAGATGAGCAAGCTGTTC-----GGCGACAGCGCGCGCTCGGAA 1574
QY      800 Ala 800
Db      1575 TCA 1577

RESULT 9
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

Alignment Scores:
Pred. No.:      2.67e-40      Length:      1914
Score:          491.50      Matches:      172
Percent Similarity: 35.36%      Conservative: 107
Best Local Similarity: 21.80%      Mismatches: 201
Query Match:     12.06%      Indels:      309
DB:              4          Gaps:      18

US-10-051-909-32 (1-800) x US-09-291-922-19 (1-1914)
QY      41 SerSerGlnGluProValThr-----SerAspAspIle-----LeuGluAspLys 55
Db      35 TCAAGAAGTAGCCGTTAAAGATGCTTCCGACGAGCTCGCAAGGCGCTGAGCCAGG 94
QY      56 MetSerGly-----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeu 71
Db      95 AAGAAGGCAACGTCAGATATGCTCCATATATGTCATCTGCGCTCCATGAGCTCTGTC 154
QY      72 LeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPhe 91
```


Db 155 ATCTTGCTATGACATTGGGGTATGATGAGCGGCCATGTACATCAAGAGGACCTG 214
QY 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111
Db 215 AATATC---ACGAGCTGCAGCTGAGATCCTGATCGGATCCTCAGTCTCTAC----- 265
QY 112 ThrIleValThrThrPheSerGlyPro---LeuSerAspSerIleGlyArgArgProMet 130
Db 266 TCGCTGTTCGATCCTTCGCTGCGCGCGGACGTCCGACAGATCGGGCGCGCTTGACC 325
QY 131 LeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsn 150
Db 326 GTGCTGTTCGCCGCTGTCATCTTCTGCGGCTCTGCTCATGGGTTTCGCCGTCAAC 385
QY 151 ValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170
Db 386 TAGGCATGCTCATGGCGGGCGCTTCGTGCGCGGAGTCCGTGTGGGCTACGGGGGCATG 445
QY 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190
Db 446 ATGGCGCCGTGTACACAGCGCCGAGATTCGCTGCGCGCTCCCGTGGCTTCCTGACCACC 505
QY 191 LeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMet 210
Db 506 TTCCCGAGAGTGTTCATCAACATCCGCATCTGCTGGCTACCTGTCCAACCTTCGGCTTC 565
QY 211 SerLeuSerPro---SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSer 229
Db 566 GCGCGCTCCCGCTCCACCTCGGCTGCGCGCTCATGTCCGCATTTGGCGCAGTCCGCTCC 625
QY 230 LeuPhePhePheGlyLeuThrIlePheTyrIleuProGluSerProArgTrpLeuValSer 249
Db 626 ---GGCTGTCTCGCGCTCCTGTGTTCATGATGCCGAGTCGCTCGGTGGTGTCTTG 682
QY 250 LysGlyArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspAspVal 269
Db 683 AAGGCGCGCTCGCGAGCCGAGGCTGTGCTAGAGAAG----- 721
QY 270 SerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyAspThrSerIleGlu 289
Db 721 ----- 721
QY 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLys 309
Db 721 ----- 721
QY 310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329
Db 722 -----ACCTCTGCCACGCCAGAGAGCGCGCGAGCGGCTGGCC----- 760
QY 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349
Db 760 ----- 760
QY 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369
Db 760 ----- 760
QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
Db 760 ----- 760
QY 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409
Db 760 ----- 760
QY 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429
Db 760 ----- 760
QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449
Db 760 ----- 760

QY 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSer 469
Db 760 ----- 760
QY 470 ThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn 489
Db 761 ---GACATCAAGCGCGG----- 775
QY 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509
Db 776 -----GCGGGGATTCGGAAGGC 793
QY 510 ArgArgGlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGlu 525
Db 794 CTCGACGGGAGCGTAGTCAACCGTACCCGGCAAGAGCAAGCGCGGTGAGTTG----- 847
QY 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
Db 847 ----- 847
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
Db 847 ----- 847
QY 566 GlySerArgTrpLysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeu 582
Db 848 ---CAGTGTGAAGAAGCTCATCTGTCCCGGACCCCGGCTGTCCGACGATACTGCTC 904
QY 583 ValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyr 602
Db 905 TCGGCCGTGGTCTCACTTCTTCAGCAGAGCTTCTGGCAGCAGCTCCGTCCTCAAGTAC 964
QY 603 ThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer 622
Db 965 AGCGCCCGCTGTCAAGAGCGCGGGGATCACCGACACAAAGCTCTCGGCGCTCACAC 1024
QY 623 SerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle 642
Db 1025 TCGCGG-----GTGGCGGTGACCAAGAGCTTCTTCATCTG----- 1060
QY 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662
Db 1061 ---GTGGCACGTTCTCTGCTGACCGCGCGGCGCTGCTGTGCTGATCAGCAGC 1117
QY 663 ProIleLeuIleAlaSerLeuValIleLeu-----ValValSerAsnLeu 677
Db 1118 GCGCGGATGATGTCTCGCTCATCTGCTCGGGGTCCGGGCTCACCGTCGGGGCATCAC 1177
QY 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
Db 1178 CCGGACACCAAGTCCGCGTGGCGCGCTCGCCCTGTGCATCGCGTCAACCTGTCTACATC 1237
QY 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717
Db 1238 GCCTTCTTCTTCATCGGCTCGGGCCCATCACGGGCGGTACACCTCGGAATAATTCCCG 1297
QY 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
Db 1298 CTGCAAGTGGCGCGCTGGGCTTTCGCGGTGGTGGCGAGCAACCGGCTCACCGCGCC 1357
QY 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
Db 1358 GTCATCTCCAGTACTTCTGTCTCCCTCCAGAGCCATCACCATCGGCGGCAAGCTTCTTC 1417
QY 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777
Db 1418 CTCTACTCCGGCATCGCGCGGTGGGTTTCTTCTTCAAGTCTCCCGAGAGACA 1477
QY 778 LysGlyMetProLeuGluValIleThrGluPhePhe-----Ala 790
Db 1478 CGCGCGCGGACGCTGAGGAGATGGGCAAGCTGTTCGGCATGCCAGACAGGGCATGGCT 1537

QY 791 ValGlyAlaLysGlnAlaAlaLys 799
Db 1538 GAAGAAGCAGAAGACCGCAGCCAAAG 1564

RESULT 10

US-09-291-922-27
; Sequence 27, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-27

Alignment Scores:
Pred. No.: 4.65e-40 Length: 1872
Score: 489.00 Matches: 165
Percent Similarity: 34.22% Conservative: 106
Best Local Similarity: 20.83% Mismatches: 215
Query Match: 12.00% Indels: 306
DB: 4 Gaps: 14

US-10-051-909-32 (1-800) x US-09-291-922-27 (1-1872)

QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCysSerSerGln 43
Db 119 ATGGCTTCTGCTCGCTCCCGGAGCCGGGGGCGAGTCCATCCAAGAAC----- 166
QY 44 GluProValThrSerAspIleLeuGluAspLysMetSerGlyAlaValLeuValAla 63
Db 167 -----AAGGGCAATTTCAGTACGCGCTTCACCTGCGCC 199
QY 64 IleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAla 83
Db 200 CTCTGTGCTTCATGAGCCACCATCGTCTCGCTACGACGTTGGGTGATGAGCGTGC 259
QY 84 ValLeuTyrIleLysGluPheGlnLeuGlnAsnGlu-----ProThrValGluGly 101
Db 260 TCGCTGTACATCAAGAGGACCTGCAGATCACGACGACGCTGAGATCATGATGGCC 319
QY 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeu 121
Db 320 ATCCTGAGCGGTACCGCTCATCGGGCTTCCTC-----GGCGGAGGAGC 367
QY 122 SerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSer 141
Db 368 TCCGACTGGGTGGCGCGCGCTCACCGCTGCTTCGGCGCGCCCATCTTCAACAACGGC 427
QY 142 GlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAsp 161
Db 428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGTGCGCGCTTCGACCC 487
QY 162 GlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro 181
Db 488 GGAATCGGCGGTGGCTACGCCATCATGTGCGCGCAGTACACGCCGAGGTGTCCCG 547
QY 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPhe 201
Db 548 GCGTCGGCCCGCGGCTTCCTCATCGTCTTTCACCGAGGTTCATCATGTGGGCTCCTC 607

QY 202 LeuSerTyrCysMetValPheGlyMetSerLeuSerPro--SerProAspTrpArgIle 220
Db 608 CTGCTGCTACTCTCCAACTACGCTTCGCGCGCTCCCGCTCCACTCAGCTGCGCGCTC 667
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeu 240
Db 668 ATGCTCGCAGTCGCGCGCGCTCCCGCTCCGCG--CTGCTGCGCTCATGTGTTCGCGCATG 724
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeu 260
Db 725 CCGAGTCTCTCGCTGCTGCTCATGAAGCCGCTCGCGGACGCGCGCGCTTCTG 784
QY 261 GlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
Db 784 ----- 784
QY 281 GluValGlyLysPThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 784 ----- 784
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
Db 784 ----- 784
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 784 ----- 784
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 784 ----- 784
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetLysSerThrLeu 380
Db 784 ----- 784
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
Db 785 -----GCCAAGACCTCCGACACG 802
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyLysAsp 420
Db 803 CCGAGAGAGCGCGTGAAGCGCTTGACCATCAAGGCTGC----- 844
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
Db 844 ----- 844
QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
Db 844 ----- 844
QY 461 GlyGluGlyLysAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480
Db 844 ----- 844
QY 481 LysTrpSerGluLysGluGlyLysAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 844 ----- 844
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 845 -----GCCGCAATCCCTAAGGAACTTACGCGACGCGAGCTGTCTCATGCTTAAG-- 892
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 892 ----- 892
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 893 -----ACAAAGCGCGCCAGAGACAGCGTG----- 919
QY 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPhe-----GluProGlyVal 577

Db 920 -----TGGAAGAGCTCATCTTTTCGCCGACCCCGCATG 955
QY 578 ArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleAsn 597
Db 956 CGGGCATACTGCTCTCGCGCGCTCGGCATCTTCTTTCAGCAGCGGCGCTCCGAC 1015
QY 598 GlyValLeuTyrThrProGlnIleLeuGlnIleAlaGlyValAlaValIleLeuSer 617
Db 1016 TCCGTCTGCTCTATAGCCACGCGTGTCCAGAGCGCGGCATCACCGGCACAAACAC 1075
QY 618 LysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeu 637
Db 1076 CTGCTCGGC-----GCCACATGCGGCATGGGGTTCATGAAGACGCTTTC 1120
QY 638 MetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeu 657
Db 1121 ATCCTG-----GTGGCCAGTTCACAGCTCAACCGCTCGGACGCGCGCTG 1168
QY 658 LeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal-----IleLeu 672
Db 1169 CTGCTGACCAAGCAGCGCGGCATGCTCGCTGTCTCATCGGCTCGGACGCGGCTCAC 1228
QY 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
Db 1229 GTCTGGGTGGCACCACCGGACGCCAAGGTCCCGTGGGCATCGGCTGTGCATCGTTC 1288
QY 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 1289 ATCTGGCTTACGTGTCTTCTTCTTCATCGGCTCGGCGCTCACACGCGTTCACAC 1348
QY 713 AlaGlnIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db 1349 TCGAGGTCTTCCACATCGCGGTGGCGCGCTGGCTTCGCGCTGGCACGTATGCAAC 1408
QY 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 1409 CGCTGACCAAGCGCGCGGTCTTCATGCTCTTCTTCCAAAGGCATCACCATC 1468
QY 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db 1469 GCGGCGAGCTTCTTCTGTACGCGGCATCGCGCGCATAGGATGGAATTTCTTCTTACC 1528
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe----- 789
Db 1529 TTCATTCGAGACGCGTGGCTGCGCTCGAGAGATAGGAACTTTTCGCGCATGACG 1588
QY 790 -----AlaValGlyAlaIleGlnAlaAlaLys 799
Db 1589 GACACGCGCGCTGGAAGCCCAAGACACCGCCACGAAA 1624
RESULT 11
US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER APPLICATION NUMBER: 1999-04-14
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-25

Alignment Scores:
Pred. No.: 2.34e-37 Length: 2089
Score: 463.50 Matches: 166
Percent Similarity: 35.82% Conservative: 108
Best Local Similarity: 21.70% Mismatches: 200
Query Match: 11.37% Indels: 291
DB: 4 Gaps: 17
US-10-051-909-32 (1-800) x US-09-291-922-25 (1-2089)
QY 52 LeuGluAspLysMetSerGly-----AlaValLeuValAlaIleValAlaSer 67
Db 104 GTCGAGCCCAAGAGAAGGCAACGTGAGTTCCCTTCGCCCTGCGCATCTCGCTCC 163
QY 68 IleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIleAlaAlaValLeuTyrIle 87
Db 164 ATGACCTCCATCTCTCTCGGCTACGACATCGGCGTGATGAGCGGAGCGCTGTACATC 223
QY 88 LysLysGluPheGlnLeuGlnAsnGluProThrValGlu-----GlyLeuIleVal 104
Db 224 CAGAAGATCTGAAGATC--AACGACACCAAGCTGAGGTCCTCATGGGCATCTCAAC 280
QY 105 SerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSer 124
Db 281 GTGTACTCGCTCATGTGCTCTTCGCG-----GCGGGCGGACGTCCGACTGG 328
QY 125 IleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIle 144
Db 329 ATCGCGCGGCTTCAACCATCGTCTTCGCGCGCTCATCTTTCGCGCGCGCTCATC 388
QY 145 MetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGly 164
Db 389 ATGGGCTTCTCCGTAACATAAGCCATGTCGCGCGCTTCGCGCGCGCATCGGC 448
QY 165 IleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIle 184
Db 449 GTGGGTACGCTTCATGATCGCGCGCGGTGAACACGGCGGAGGTGTCCCGCTCTGCC 508
QY 185 ArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyr 204
Db 509 CGTGGGTTCATCATCTCTTCCCGAGGTTTCATCAACTTCGCGCATCTCTCGGATAT 568
QY 205 CysMetValPheGlyMetSer--LeuSerProSerProAspTyrArgIleMetLeuGly 223
Db 569 GTCTCAACTTCGCTTCCCGCGCTCTCTCCCTCGCGCTGCGCATATGCTCGGC 628
QY 224 ValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSer 243
Db 629 ATAGCGCGGTGCGCTCCGTC--CTGCTCGGTTTCATGCTGCTCGGCATGCGGAGTCT 685
QY 244 ProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeu 263
Db 686 CCGCGGTGCTCTCATGAAGGCGCTGCGGACGCCAAGGTTGTGCTTGCCAAAGACG 745
QY 264 ArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuValGly 283
Db 745 ----- 745
QY 284 GlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAspAspLeu 303
Db 746 TCCGACACG----- 754
QY 304 ValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlnSerTrp 323
Db 755 -----CCGAAGAGCGCGCGGAGCGC 775
QY 324 IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArg 343
Db 776 ATCGC----- 781
QY 344 HisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGly 363


```
Db 781 ----- 781
QY 364 SerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsn 383
Db 781 ----- 781
QY 384 PheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGlnGlnTrpAspGlnGlu 403
Db 781 ----- 781
QY 404 AsnLeuHisArgAspArgGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp 423
Db 782 ----- 799
      GACATTAAGACTGCCGCC-----
QY 424 AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGlyGlyLysAspIleVal 443
Db 799 ----- 799
QY 444 HisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGlyGly 463
Db 799 ----- 799
QY 464 GlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpIleTrpSer 483
Db 799 ----- 799
QY 484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGln 503
Db 799 ----- 799
QY 504 GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspVal 523
Db 800 ---GGCATCCCTCTGGGCGCTCGACGGCGACGCTGTCGCCCGTCCCAAA----- 844
QY 524 PheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSer 543
Db 845 -----AAC 847
QY 544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563
Db 848 AAAGGAAGCAGCAGAGAGAGCCGTTTGAAGACCTCATCTGTACCGCACCATAGCC 907
QY 564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuVal 583
Db 908 -----ATGCCGCACATCCTCATCGCG 928
QY 584 GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTrpThr 603
Db 929 GGAATCGGCATCCACTTCTTCAGCAGAGTCTTCGGGCAATGACGCCGCTGCTTAACAGC 988
QY 604 ProGlnIleLeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623
Db 989 CCGCTAGTTTCAAGAGCGCGGCATCAGC-----GGCAGACGCCGT 1030
QY 624 ---AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle 642
Db 1031 CTCGCGCGCACACCGTGGGGGTTCGGGCCACCAATACGGTCTTCATCCTG----- 1081
QY 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIle 662
Db 1082 ---GTGGCCACCTTCCTCTCGACCGCATCCGCCGCCGCGCTGTGTGACCGACAGC 1138
QY 663 ProIleLeuIleAlaSerLeuVal-----IleLeuValValSerAsnLeu 677
Db 1139 GCGCGCATGCTCGTCTCTTAAGTGGGCTCGCGACGGGGCTCACCGTCATCAGCCGCCAC 1198
QY 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
Db 1199 CCGAGCAGAAAGATCACCTGGGCCATCGTCTGTGCATCTTGTGCATCATGCGCTACGTG 1258
QY 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717
Db 1259 GCCTTCTTCTCATCGGCTCGGCCCATCACGTGGGTGTACAGCTCGAGATCTTCCCG 1318
```

```
QY 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
Db 1319 CTGCAcGTGCGCGCGCTGGGCTGCTCCCTGGCGGTGGCCGTCAACCGCTGACGCGGC 1378
QY 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
Db 1379 GTGATCTCCATGACTTTCATTTCGTGCTGTCCAGGCCATGACCATCGCGCGCTTCTTC 1438
QY 758 IleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777
Db 1439 CTCTTCGCGCGCATCGCTCATTCGTCATGGGTGTCTTCTTCCTGCTTACCTGCGGAGACC 1498
QY 778 LysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla-----LysGln 795
Db 1499 CGCGCGCGCAGCGTGGAGGACATGAGCTCGTGTTCGGCAACACGCGCACAGCAGCAG 1558
QY 796 AlaAlaAlaLysAla 800
Db 1559 GCGCGCGCGGAGGCC 1573

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-2007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 2,59e-26 Length: 4411529
Score: 412.00 Matches: 162
Percent Similarity: 33.46% Conservative: 111
Best Local Similarity: 19.85% Mismatches: 204
Query Match: 10.11% Indels: 340
DB: 3 Gaps: 17

US-10-051-909-32 (1-800) x US-09-103-840A-1 (1-4411529)
QY 2 ArgSerGlySerTrpLeuAlaValGlnThrPro-----PheThrPro 15
Db 3717093 CGCTGTGGCAGCCTTCATCGAATGACTACAGCCCTATACAGACGCGGGTTCACGCC 3717152
QY 16 -----AspLeuAspArgGluArgLeuLeuProSerValValLeuAla 30
Db 3717153 GCCCGGGCGCCGACGACCCCGACCGC-----GTGTGACAGGTG 3717191
QY 31 IeuProGlyProLeuPro-----ProAlaSerCysSerSerGlnIleProValThr 47
Db 3717192 CTGAGCGCCCGGTACCGGTCAACTGACCA----- 3717221
QY 48 SerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSer 67
Db 3717222 -----GATCGGGCGCGCGCG-CTCCTGTGCGGCTCACCGCGCC 3717262
QY 68 IleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIle 87
Db 3717263 AGCGTCGCGCTCTTACGGGTACGACCTTTCGCCCATCGCGGTGCTGTCTCTC 3717322
```

```
QY      88  LybLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSer 107
Db      3717323  ACGGAGGAATTGCAACTCACCCTCGAGAA-----CAGGAGTTGCTGACCACCGGCGG 3717376
QY      108  LeuIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArg 127
Db      3717377  GTGCTCGGCCAGATCGCCGGGGCGCTTGGCGGGCGGCATCTTCCCAACGCGATCGGAGCC 3717436
QY      128  ArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrp 147
Db      3717437  AAGAAATCGGTGTGCTCATCGTCCGCGCTACGACAGTGTTCCTGCTCGGCGGAGCC 3717496
QY      148  SerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeu 167
Db      3717497  TCGGTGTCCGTCACCGATCGTGTGGTGGCGCGTCTGCTGTGGTGTGACAATCGGCGTG 3717556
QY      168  AlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeu 187
Db      3717557  TCGGTGTGTGTGTGTGCGGTGTATGTGGCCGAGTCGCGCGCGCGCGGTGCGTGGGTGG 3717616
QY      188  LeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetVal 207
Db      3717617  TTGGTGACCGCGTATCAGCTGGCGACGCTTAGCGGATATCGTCTCGGTAC-----CTG 3717670
QY      208  PheGlyMetSerLeuSerProSerProAspTyrArgIleMetLeuGlyValLeuAlaIle 227
Db      3717671  GTGCGCTACCTGTGTGGCCGATCGACGCGCTGGCGCGATGTTGGGGCTGGCGCGCGCG 3717730
QY      228  ProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeu 247
Db      3717731  CCGGCCACGCTGCTGTGTCGGCTG---TTGTGGCGCATGCGCCGATACCGCCCGCTGGTAT 3717787
QY      248  ValSerLysGlyArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAsp 267
Db      3717788  CTGCTCAAGGGCCGATCGCCGACGCGCGTACGCGCTGCGCGCGATCCAGCCGAGGCC 3717847
QY      268  AspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSer 287
Db      3717848  GACATTCGATCGCGAGCTGGCC----- 3717868
QY      288  IleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGly 307
Db      3717868  ----- 3717868
QY      308  AspLysGluGlnIleThrLeuTyrGlyProGluGlnGlnSerTrpIleAlaArgPro 327
Db      3717868  ----- 3717868
QY      328  SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db      3717868  ----- 3717868
QY      348  ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367
Db      3717868  ----- 3717868
QY      368  AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db      3717868  ----- 3717868
QY      388  PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
Db      3717868  ----- 3717868
QY      408  AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427
Db      3717869  ---GATATGGCGCGCGGTGACGACGACGCGCGCGGTATCGCGGAATGCTGCGGCGG 3717925
QY      428  ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447
Db      3717926  CCGTATCTG----- 3717934
```

```
QY      448  ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyLysAspGlyVal 467
Db      3717934  ----- 3717934
QY      468  SerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGly 487
Db      3717934  ----- 3717934
QY      488  GluAsnGlyArgGlyGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValPro 507
Db      3717934  ----- 3717934
QY      508  GlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySer 527
Db      3717934  ----- 3717934
QY      528  GluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAla 547
Db      3717934  ----- 3717934
QY      548  GluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySer 567
Db      3717934  ----- 3717934
QY      568  ArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIle 587
Db      3717935  -----CGGGCCACGCTGTTCGTATCGCGCTC 3717961
QY      588  GlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeu 607
Db      3717962  GGCTTCCTCGTCCAGATCACCGGATCAACGGCATCATCTACAGTCCGCACTTTC 3718021
QY      608  GluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSer 627
Db      3718022  GCCCGCATGGGCTTCGCG-----GGCTATTTCGGATGCTTGCCCTGCGCGCATG 3718072
QY      628  IleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeu 647
Db      3718073  GTGCAAGTCGCGGCTTGGCGCGGTG-----TGTGCC-----TCGCTGTTT 3718114
QY      648  LeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAla 667
Db      3718115  CTGCTGATCGGCTGGCGCGTCCGCGCATCTGTTCGCGCATCGCGAGATGATCACCC 3718174
QY      668  -----SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
Db      3718175  GCAGATGCCGTGCTGATCACCGTATTCCGCCAACGACTCCGATGTGGCACG-----GGG 3718228
QY      686  AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
Db      3718229  CTGCTGTGGGTTGCGCGCGGTGCTGTTCATCATCGGTTCACTTCGGATTCCGGATTCGGC 3718288
QY      706  ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu---Cys 724
Db      3718289  TCGCTGTCTGGTGTACCGCGGAGAGCTTCCCGCGCTGCGGTGATGGGATCG 3718348
QY      725  IleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuPro 744
Db      3718349  AGCCCGATGCTCACTCGACACTGACGCGCAACGCGATCGTTGCCCTCTCGCTC--- 3718405
QY      745  ValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeu 764
Db      3718406  ACCATGCTGCGTGTGCTCGCGCGGCGCAGCGCTTTTCGGGTCTTCGGACGTTCCCGCTC 3718465
QY      765  IleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluVal 784
Db      3718466  GTGCGTTCGTGTGTGTACCGCTTTCGCGCGGAGACCAAGGCGCGCAACTCGAGGAG 3718525
QY      785  IleThrGluPhePheAlaValAlaLysGlnAlaAlaLysAla 800
Db      3718526  ATCCGGCACTTCTGGAGAACGCGCGCGCTGGCCCGCGAGCGGTCA 3718573
```

RESULT 13


```

QY      547  aGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySe 567
Db      3715494  -----
QY      567  rArgTrpLysAspLeuPheGluProGlyValAlaArgAlaLeuLeuValGlyValGlyI 587
Db      3715495  -----
QY      587  eGlnIleuGlnGlnPheAlaGlyIleasnGlyValLeuTyrTyrThrProGlnIleLe 607
Db      3715521  CGGCTTCCCTCGTCCAGATCACCGGATCAACGCGATCATCTACTACAGTCCGCGACTTTT 3715580
QY      607  uGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlase 627
Db      3715581  CGCCGCCATGGGCTTCGCG-----GGCTATTTCGCGATGCTTCGCCCTGCCCCGCGAT 3715631
QY      627  rIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLe 647
Db      3715632  GGTGCAAGTCGCGCGCTTGCGCGCGGTG-----TGTGCC-----TCGCTGTT 3715673
QY      647  uLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAl 667
Db      3715674  TCTGTGATCGGCTGGGCCGTCGCCCGATCTGTTGTCCGCGATCGCGACGATGATCAC 3715733
QY      667  a-----SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHi 685
Db      3715734  CGGAGATGCCCGTGTGATCACCGGTATTCCGCCAACGACTCCGATGTGTGCACG-----GG 3715787
QY      685  sAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheG 705
Db      3715788  GCTGTGTGGGTTGCGCCGCGCTGCTGTTCATCATCGGGTTCAACTTCGGATTCCG 3715847
QY      705  yProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysI 725
Db      3715848  CTCGCTGTGTGGGTGTACGCGCGAGAGCTTCCGTCGCGTGGTCGATGGGATC 3715907
QY      725  eAlaIle--CysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuPr 744
Db      3715908  GAGCCTGATGCTCACCTCGACACTGACGCGCAACGCGATCGTTCGCCCTTCTCGCTC-- 3715965
QY      744  oValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLe 764
Db      3715966  -ACCATGCTCGCTGTGCTCGCGCGCGCAGGCGCTTTTCGCGCTTCGGCACGTTGCCCT 3716024
QY      764  uIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluVa 784
Db      3716025  CGTCGCGTTGCTGTGTGTACCGCTTTCGCCCGGAGACCAAGGCGCAACTCGAGGA 3716084
QY      784  lIleThrGluPhePheAlaValAlaGlyAlaLysGlnAlaAlaAlaLysAla 800
Db      3716085  GATCCGCGCACTTCTGGAGAACGCGCGCGCTTGCGCCCGCGAGCGGTCA 3716133

RESULT 14
US-09-643-597-135
; Sequence 135, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597

```

```

; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-135

Alignment Scores:
Pred. No.: 5 67e-26 Length: 2856
Score: 355.00 Matches: 164
Percent Similarity: 33.29% Conservative: 107
Best Local Similarity: 20.15% Mismatches: 206
Query Match: 8.71% Indels: 337
DB: 4 Gaps: 23

US-10-051-909-32 (1-800) x US-09-643-597-135 (1-2856)

QY 18 AsparGargGluaRgLeuLeuProSerValValLeuAlaLeuProGly----- 33
Db 86 GACCGAGCAGCAGCCTGAGCGGAGAGCGCGCTCGACGCCCTCGCCACCCGCTAC 145
QY 34 -----ProLeuProProAla-SerCysSerSerGlnGluProValThrSerAsp 50
Db 146 CCGGCGCAGCCAGAGCCAGCCAGCGCAGC---GTCGATGAGGCCAGCAAG----- 197
QY 50 pIleLeuGluAspLysMetSerGlyAlaValLeuValAlaIle---ValAlaSerIleG 69
Db 198 -----AAGCTGACGGGTGCGCTCATGCTGCTGTGGAGAGACAGTGTGG 244
QY 69 yAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLys 89
Db 245 CTCCTGACGTTGGCTACACACTGAGTCATCATGCCCCAGAGGTGATC--GA 301
QY 89 sGluPheGlnLeuGln-----AsnG 96
Db 302 GGAGTTCTACACCAAGACATGAGTCCACCGCTATGGGAGAGACATCTGCCACCA 361
QY 96 uProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValTh 116
Db 362 CACCACGCTGTGTCCTCTCAGTGGCCATCTTTCTGTTGGGGCATGATGGCTCT 421
QY 116 rPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSer 136
Db 422 CTCTGTGGCCTTTTCGTTAACCGCTTTGGCCGCGGAATTCATATGATGAACCT 481
QY 136 eLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer-----ProAsnValTyr 153
Db 482 GCTGGCCTTCGTGTCCGCCGTGCTCATGGGCTTCGAAACTGGGCAAGTCTTGA 541
QY 153 lLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeu 173
Db 542 GCTGATCCTGGGCGCTTCATCATCGGTGTACTGCGGCTGACCAAGCTTCGTGC 601
QY 173 oLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeu 193
Db 602 CATGATATGTGGGTGAAGTGCACCCACAGCCTTTCGTGGGCGCTGGGCACCTG 661
QY 193 nPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet-- 212
Db 662 GCTGGGCATCGTGTGCGCATCCTCATGCGC-----CAGGTGTGCGCTGAC 715
QY 212 userProSerProasp--TrpArgIleMetLeuGlyValLeuAlaIleProSerLe 231
Db 716 CATGGGCAACAGACCTGTGGCCCCGTGTGATGAGCATCATCTTCAATCCGCGCT 775
QY 231 ePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrp--LeuValSer 250
Db 776 GCAGTGC--ATCGTGTGCGCTTCTGCCCCGAGAGTCCCCGCTTCTCTCATCA 832
QY 250 sGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspVal 270

```

```
Db      833 CAACGAGAGAACCGGCCAAGAGTGTGCTAAAGAGCTGCGCGGACAGCTGACGTGAC 892
QY      270 rglyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluG1 290
      : : : : |
Db      893 CCATGACCTG----- 902
QY      290 uTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysG1 310
      : : : : |
Db      902 ----- 902
QY      310 uGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysG1 330
      : : : : |
Db      902 ----- 902
QY      330 yProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnG1 350
      : : : : |
Db      902 ----- 902
QY      350 nSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPr 370
      : : : : |
Db      902 ----- 902
QY      370 oGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVa 390
      : : : : |
Db      902 ----- 902
QY      390 lThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspG1 410
      : : : : |
Db      903 -----CAGGAGATGAAGGAGAGAG----- 920
QY      410 uGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlyAspAsnLeuHisSerProLeuLe 430
      : : : : |
Db      920 ----- 920
QY      430 uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySe 450
      : : : : |
Db      921 -AGTCGGCAG----- 929
QY      450 rAlaLeuSerMetArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerTh 470
      : : : : |
Db      930 -ATGATGCGGAGAGAGAGAGTCACCATCTCTG----- 959
QY      470 rAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnG1 490
      : : : : |
Db      959 ----- 959
QY      490 yArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerAr 510
      : : : : |
Db      959 ----- 959
QY      510 gArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheVa 530
      : : : : |
Db      959 ----- 959
QY      530 lHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProAr 550
      : : : : |
Db      959 ----- 959
QY      550 gMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLy 570
      : : : : |
Db      959 ----- 959
QY      570 sAspLeuPhe---GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnI1 589
      : : : : |
Db      960 -GAGCTGTTCCGCTCCCGCCTACCGCCAGCCCATCTCATCGCTGTGCTGCTGCAGCT 1018
      : : : : |
QY      589 eLeuGlnGlnPheAlaGlyIleAsnGlyValLeuLeuTyrTyrThrProGlnIleLeuGluG1 609
      : : : : |
Db      1019 GTCCACAGCAGCTGTCTGGCATCAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAA 1078
      : : : : |
QY      609 nAlaGlyVal---AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerI1 628
      : : : : |
Db      1079 GCGCGGGGTGCAGCAGCCTGTGTATGCACCATTTGCTCCGGTATCGTCAACACGCGCTT 1138
```

```
QY      628 eLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLe 648
      : : : : |
Db      1139 CACTGTCGTGCTGCTG-----TTTGT 1159
QY      648 uMetAspLeuSerGlyArgArgPheLeu---LeuLeuGlyThrIleProIleLeu---Il 666
      : : : : |
Db      1160 GGTGAGCGAGCAGCGCCGCGGACCTGACCTCATAGGCTCGCTGGCATGGCGGTTG 1219
      : : : : |
QY      666 eAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAl 686
      : : : : |
Db      1220 TGCCATACTCATGACCATCGCGCTAGCAGCTGAGCAGCTACCTGAGTGTCTTAT-- 1277
      : : : : |
QY      686 aLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPr 706
      : : : : |
Db      1278 ----CTGAGCATGCTGGCCATCTTGCGCTTGTGGCCTTCTTGAAGTGGGTCTGGCCC 1333
      : : : : |
QY      706 oIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAl 726
      : : : : |
Db      1334 CATCCCATGGTTCATCGTGGCTGAACCTTTCAGCCAGGGTCCACGCTGCCATTGC 1393
      : : : : |
QY      726 aIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMe 746
      : : : : |
Db      1394 CGTTGACGGCTTCTCCAACTGACCTCAAAATTTCATTGTGGCATGTGCTCCAGTATGT 1453
      : : : : |
QY      746 tLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSe 766
      : : : : |
Db      1454 GGAGCAACTGTGTGT--CCCTACGTCTTCATCATCTTCACTGTGTCTGTCTGT 1510
      : : : : |
QY      766 rPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleTh 786
      : : : : |
Db      1511 CTTCACTTCACTCACTTCAAGTTCCTGAAGCTAAAGCCGACCTTCATGAGATCGC 1570
      : : : : |
QY      786 rGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
      : : : : |
Db      1571 TTCCGGCTTCCGGCAGGGGGAGCCAGCCCAAGTATAAG 1610
      : : : : |
```

```
RESULT 15
US-09-480-884A-135
; Sequence 135, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Panger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-135
```

```
Alignment Scores:
Pred. No.: 5.67e-26 Length: 2856
Score: 355.00 Matches: 164
Percent Similarity: 33.29% Conservative: 107
Best Local Similarity: 20.15% Mismatches: 206
Query Match: 8.71% Indels: 337
DB: 4 Gaps: 23
```

```
US-10-051-909-32 (1-800) x US-09-480-884A-135 (1-2856)
QY      18 AspArgArgGluArgLeuLeuProSerValValLeuAlaLeuProGly----- 33
      : : : : |
Db      86 GACCGAGACAGACGCTTGAAGCGGAGAGCGCGCTCGCACGCCCGTGCACCCGCGTAC 145
      : : : : |
```

QY 34 -----ProleuProProAla-SerCysSerSerGlnGluProValThrSerAspAs 50
Db 146 CCGGGCCAGCCAGAGCCAGCCAGCCAGC---GCTGCCATGGAGCCAGCAGCAG--- 197
QY 50 pIleuGluAspLysMetSerGlyAlaValLeuValAlaIle---ValAlaSerIleG1 69
Db 198 -----AAGCTGACGGGTCCCTCATGCTGGCTGGGAGAGCAGAGCTTGG 244
QY 69 yAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLy 89
Db 245 CTCCTGCAGTTTGGCTACACACTGGAGTCAATGATGCCCCAGAGAGGTATC--GA 301
QY 89 sGluPheGlnLeuGln-----Aang1 96
Db 302 GGAGTTCTACAAACAGACATGGGTCCACCGCTATGAGAGAGCATCTGCCACACGCT 361
QY 96 uProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrTh 116
Db 362 CACGACGCTCTGGTCCCTCTCAGTGGCCATCTTCTCTGGGGGCATGATGGCTCCTT 421
QY 116 rPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer11 136
Db 422 CTCTGTGGGCTTTTCGTTAACCGCTTTGGCCGGCGAATTCATGCTGATGATGAACCT 481
QY 136 eleuTyrPhePheSerGlyLeuIleMetLeuTrpSer-----ProAsnValTyrVa 153
Db 482 GCTGGCTTCTGTCCTCCCGCTCATGGGCTTCTGAAACTGGGCAAGTCTTGTGAGAT 541
QY 153 lLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPr 173
Db 542 GCTGATCTGGGGCCGCTTCATATCGGTGTGTACTGCGGCTGACACAGGCTTCGTGCC 601
QY 173 oleuTyrIleSerGlnIleAlaProSerGlnIleArgGlyLeuLeuAsnThrLeuProG1 193
Db 602 CATGTATGTGGTGAAGTGTACCCACAGCCTTTCGTGGGCGCTGGCACCCTGCACCA 661
QY 193 nPheSerGlySerGlyLysMetPheLeuSerTyrCysMetValPheGlyMet---SerLe 212
Db 662 GCTGGGCATCTGTCGGCATCTCATCGCC---CAGGTGTTGGCGCTGACCTCCAT 715
QY 212 uSerProSerProAsp---TrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPh 231
Db 716 CATGGGCAACAAGACCTGTGGCCCCCTGCTGTGAGCATCATCTTCATCCCGGCCCTGCT 775
QY 231 ephePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrp--LeuValSerLy 250
Db 776 GCAGTGC--ATCGTGTGCCCTTCTGCCCGAGAGTCCCGCTTCTGCTCATCAACG 832
QY 250 sGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValse 270
Db 833 CAACGAGAGAACCGGCGCAAGTGTGTTAAAGAAAGCTGCGCGGACAGCTGACGTGAC 892
QY 270 rGlyGluLeuSerLeuLeuGlnGlyLeuGluValGlyGlyAspThrSerIleGluG1 290
Db 893 CCAATGACCTG----- 902
QY 290 uTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysG1 310
Db 902 ----- 902
QY 310 uGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysG1 330
Db 902 ----- 902
QY 330 yProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnG1 350
Db 902 ----- 902
QY 350 nSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPr 370
Db 902 ----- 902

QY 370 oGlnAlaGlyLysMetLysSerThrLeuPheProAsnPheGlySerMetPheSerVa 390
Db 902 ----- 902
QY 390 lThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspG1 410
Db 903 -----CAGAGATGAAGGAAGAG----- 920
QY 410 uGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLe 430
Db 920 ----- 920
QY 430 uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisIleGlyHisArgGlySe 450
Db 921 -AGTCGGCAG----- 929
QY 450 rAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyLysAspGlyValSerSerTh 470
Db 930 -ATGATCGCGGAGAGAAGGTCACCATCTG----- 959
QY 470 rAspIleGlyGlyTyrPglLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnG1 490
Db 959 ----- 959
QY 490 yArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerAr 510
Db 959 ----- 959
QY 510 gArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheVa 530
Db 959 ----- 959
QY 530 lHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProAr 550
Db 959 ----- 959
QY 550 gMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLy 570
Db 959 ----- 959
QY 570 sAspLeuPhe---GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnI1 589
Db 960 -GAGCTGTCCGCTCCCGCCTACCAGCCAGCCCATCTCATCGCTGTGCTGCAGCT 1018
QY 589 eleuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluG1 609
Db 1019 GTCCACAGCAGCTGTCTGGCATCAACGCTGTCTTCTATTACTCCACGAGCATCTCGAGAA 1078
QY 609 nAlaGlyVal--AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSer11 628
Db 1079 GCGCGGGGTGCAGCAGCCTGTGTATGCCACCATTTGGCTCCGGTATCGTCAACAGGCCCTT 1138
QY 628 eleuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLe 648
Db 1139 CACTGTCTGTGCTG-----TTTGT 1159
QY 648 uMetAspLeuSerGlyArgArgPheLeu--LeuLeuGlyTyrIleProIleLeu---11 666
Db 1160 GGTGAGCGAGCAGGCGCGGAGCCTGCACCTCATAGGCTCGCTGGCATGGCGGTTG 1219
QY 666 eAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAl 686
Db 1220 TGCCATACTCATGACCATCGCGCTAGCAGCTGAGCAGCATACCTGAGATCTCTAT-- 1277
QY 686 aleuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPr 706
Db 1278 -----CTGAGCATGCTGGCCATCTTGGCTTTGTGGCCTTCTTGAAGTGGGTCTG 1333
QY 706 oIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAl 726
Db 1334 CATCCCAATGTTTCATCGTGGCTGAACCTTTCAGCCAGGGTCCACGCTCAGCTGCCATTGC 1393
QY 726 aIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMe 746


```

Db      1394 CGTTGCAGGCTTCTTCCAACTGACCTCAATTTTCATTGTGGGCATGTGCTTCCAGTATGT 1453
QY      746 tleuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSe 766
Db      1454 GGAGCAACTGTGTGT---CCCTACGTCTTTCATCATCTTCACTGTGTCTCCGTGTCTGT 1510
QY      766 rPheValPheValPheLeuIysValProGluThrIyegIyMetProLeuGluValIleTh 786
Db      1511 CTTCACTCTTCACTTCAAGTTCTGTGAGACTTAAAGGCCGACCTTCGATGAGATCGC 1570
QY      786 rGluPhePheAlaValGlyAlaIysGlnAlaAlaIys 799
Db      1571 TTCCGGCTTCCGGCAGGGGGAGCCCAAGTGATAG 1610

```

Search completed: January 6, 2004, 01:58:24
 Job time : 12167 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 18:58:54 ; Search time 543 Seconds
(without alignments)
3977.073 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPTFPDLDR.....PLEVITEFFAVGAKQAAMA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool/US10051909/runat_05012004_094741_22299/app_query.fasta_1.967
-DB=N_Geneseq_19Jun03 -QFMT=fastp -SUFFIX=mg -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909_@CGN_1_1_326_@runat_05012004_094741_22299 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3517	86.3	2824	24	ABK51962	Corn contig encodi
2	3517	86.3	2824	25	ABX93198	CDNA encoding corn
3	2674	65.6	2601	24	ABK51965	Soybean contig enc
4	2674	65.6	2601	25	ABX93201	CDNA encoding soyb
5	2584	63.4	2190	24	ABZ13510	Arabidopsis thalia
6	2275	55.8	2205	24	ABZ14449	Arabidopsis thalia
7	1670.5	41.0	1692	24	ABK51966	Soybean CDNA clone
8	1670.5	41.0	1692	25	ABX93202	CDNA encoding soyb
9	1424	34.9	1487	24	ABK51968	Wheat contig encod
10	1424	34.9	1487	25	ABX93204	CDNA encoding whea
11	943	23.1	1009	24	ABK51969	Wheat cDNA clone w
12	943	23.1	1009	25	ABX93205	CDNA encoding whea
13	587.5	14.4	1518	24	ABL41880	Nucleotide sequenc
14	562	13.8	870	24	ABK51964	Rice cDNA clone r1
15	562	13.8	870	25	ABX93200	CDNA encoding rice
16	539	13.2	1395	24	ABK73616	Bacillus lichenifo
17	520.5	12.8	1826	21	AAC42332	Arabidopsis thalia
18	496	12.2	1853	24	ABK51973	Soybean contig enc
19	496	12.2	1853	25	ABX93209	CDNA encoding soyb
20	494.5	12.1	2017	24	ABK51972	Rice contig encodi
21	494.5	12.1	2017	25	ABX93208	CDNA encoding rice
22	493.5	12.1	1527	21	AAC43261	Arabidopsis thalia
23	493	12.1	1482	24	ABZ12990	Arabidopsis thalia
24	491.5	12.1	1914	24	ABK51971	Corn cDNA clone ce
25	491.5	12.1	1914	25	ABX93207	CDNA encoding corn
26	489	12.0	1566	24	ABZ13964	Arabidopsis thalia
27	489	12.0	1872	24	ABK51975	Wheat cDNA clone w
28	488	12.0	1872	25	ABX93211	CDNA encoding whea
29	488	12.0	1925	24	ABL41874	Nucleotide sequenc
30	488	12.0	2056	24	ABL41873	Nucleotide sequenc
31	481	11.8	1766	24	ABL41872	Nucleotide sequenc
32	475	11.7	1644	24	ABZ12515	Arabidopsis thalia
33	468	11.5	1862	20	AAC24474	Spinach glucose tr
34	467	11.5	2134	21	AAC36954	Arabidopsis thalia
35	465.5	11.4	1690	24	ABL41879	Nucleotide sequenc
36	463.5	11.4	1662	24	ABL41878	Nucleotide sequenc
37	463.5	11.4	2089	24	ABK51974	Wheat cDNA clone w
38	463.5	11.4	2089	25	ABX93210	CDNA encoding whea
39	453.5	11.1	1905	22	AAH88793	Sugar transporter
40	453.5	11.1	3370	19	AAV11474	T. halophilus xylu
41	452	11.1	12354	22	AAS46243	DNA encoding novel
42	452	11.1	12354	22	AAS46261	DNA encoding novel
43	452	11.1	12354	22	AAS46267	DNA encoding novel
44	447.5	11.0	34980	22	AAH64966	C glutamicum codin
45	444	10.9	1473	22	AAH65168	C glutamicum codin

ALIGNMENTS

RESULT 1
ABK51962
ID ABK51962 standard; cDNA; 2824 BP.
XX
AC ABK51962;
XX
DT 27-AUG-2002 (first entry)
DE Corn contig encoding Arabidopsis thaliana-like sugar transport protein.
XX
XX Corn, Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant; gene; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers


```
FT CDS 238..2481
FT /*tag= a
FT /product= "Corn Arabidopsis thaliana-like sugar
FT transport protein"
XX
PN US6383776-B1.
XX
XX 07-MAY-2002.
XX
XX 14-APR-1999. 99US-0291922.
XX
XX 24-APR-1998. 98US-083044P.
XX
PA (DUPO ) DU FONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-451386/48.
DR P-PSDB; AAU97201.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 3; Column 25-28; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC representing a contig assembled from various corn cDNA clones encodes
CC an Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other;

Alignment Scores:
Pred. No.: 3.81e-298 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 24 Gaps: 2

US-10-051-909-32 (1-800) x ABK51962 (1-2824)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db ATGGGGGGCGCGCTGATGCTGCATCGCGGCTTATCGCACTTGCTGCAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleValysGluPheGlnLeuGlnAsn 95
Db GACAAATGCGACAATTGCTGAGCGCTCTGTACATTAAGAAGAAATTCAACCTGCAGAC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db GAGCCTCTGATCGAAGGCTCATCGTCGCATGTCCTCATTTGGGGCAACAGTCATCACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db ACATCTCCGGGGCCAAAGGCTGACTGCGTTGGTAGAGGCCCATGCTGTCGCTCGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db GTCTCTACTTCTGTCAGTGGCTGTGATGCTTTGGCGCCAATTGTGTACATCTTGCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db CTGCAAGGCTCATGTATGGGTTCGGTATCGGTTTGGCGGTCACTTGTCTCTCTAC 597
```

```
QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194
Db ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNTNGAACACAGTTCGCCAGTTC 657
QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db ATGGGGTCAAGGAGGATGTTCTCTCTCTACTGCATGGTGTGGATGTCCTCATG 717
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db CCCAAACCTGATTGGAGGCTCATGCTTGAGTTCTGTCCATCCCGTCACTTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIysGlyArgMet 253
Db GGACTGACTGTCTTCTACTTGCCCTGAATCACCAAGGTGGCTGTGNAGCAAGAAGATG 837
QY 254 AlaGluAlaValysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeu 273
Db CCGGAGGCGAAGAGACTGNTGCCAAGGCTGCGGGAAGAAGATGTCANGGAGANG 897
QY 274 SerLeuLeuLeuGlnGlyLeuGluValGlyIysAspThrSerIleGlnGluTyrIleIle 293
Db GCTCTTCTAGTTGAAGGTTTGGGGTTCGGTAAAGATACACGTAATTNAGAGTACATCAT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db GGACCTGCCACCGAGCAGCCGATGATCTTGTAACGCGTGATTAAGAACAAATCACA 1017
QY 314 LeuTyrGlyProGluGlnGlyGlnSerTrpIleAlaArgProSerIysGlyProIleMet 333
Db CTTTATGGCGCTGAAGAAGCCAGTCATGGAATTGCTCGACCTTCAAGGACCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db CTGGAAGTGTCTTCTCTTGCATCTGTCATGGAGACATGTTGAACCAAGTGTATCCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGlyAsnMetProGlnAlaGly 373
Db CTTATGATCCGATTGTGACACTTTTGTGTAGTGTCCATGAGAATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db GGAAGTATGAGGACACATGTTTCCAACTTGGAAATGTTGAGTATGTCACAGATCAG 1257
QY 394 HisAlaLysAsnGlnTrpAspGluGlnAsnLeuHisArgAspAspGluTyrAla 413
Db CATGCCAAAATGACAGCATGGGATGAAGAGAACTTTCATAGGATGACGAGAGTACGCA 1317
QY 414 SerAspGlyAlaGlyIysPTrpGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db TCTGATGTCGACAGAGGTGACTATGAGGACATCTCCATAGCCCATGCTGTCAGGACG 1377
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
Db GCAACAGGTCCGGAAGGGAAGGACATTGTGCAACCATGTCACCGTGAAGTCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlnGlyIysGlyIysAspGlyValSerSerThrAspIleGly 473
Db ATGAGAAAGCAAGCCTTTAGGGGAGGTGAGATGTTGTGAGCAGACATGATATCGGT 1497
QY 474 GlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyLysAsnGlyArgLysGlu 493
Db GGGGATGCGAGCTTCTTGAAATGTCAGAGAAGAAAGGTGAGAAATGTTAGAAAGGAA 1557
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGlnGlyValProGlySerArgGlySer 513
Db GGTGTTTCAAAAGAGTCTACTTGACCAAGAGGAGTTCCTGGCTCAAGAGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyIysAspValPheGluGlySerGluPheValHisAlaAla 533
Db ATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAAGGTAAGTGTATCATGCTGCT 1677
```

```

QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
DB 1678 GCTTAGTAGTACAGTACGACCTTTCTCAAGGGCTGTGCTGAACACGACATGACAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
DB 1738 GCTGCCATGTTCAACCCACTGAGTAGCTGCCAAGGTTCAAGTTGAAAGATTGTGTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePhe 593
DB 1798 GAACCTGAGTAGAGCGCTGCCCTGTAGTCGGTGTGGAATTCAGATCCTCAACAGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAla 613
DB 1858 GCTGAATTAACGGTGTCTGTAATAACCCACAATTTCTTGAGCAAGCTGTGTGCA 1917
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
DB 1918 GTTATTCCTTCCAATTTGGTCTCAGCTCGGCATCAGCATCTTGATCAGTTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
DB 1978 ACTACCTTAATGCTCTCTGATGCTTGCTTGCCATGCTGCTTATGATCTTCCGGA 2037
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
DB 2038 AGAAGTTTGTGCTGAGGACAAATCCAAATCTGATGATGATCTAGTATCTGTT 2097
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
DB 2098 GTGTCCAATCTAATGATTTGGTGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
DB 2158 ATGCTACTTCTGCTGCTGCTGTTATGGATTGGTCCCATCCCAACATTTTATGTGCA 2217
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
DB 2218 GAGATCTTCCAAACCAAGGGTGTGCTGCTCTGTATGCGCATTTGTGCTTACATCTCG 2277
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
DB 2278 ATCGAGATATCATCGTCACCTACACCTTCTGTGATGCTGAATGCTATGAGCTGGCG 2337
QY 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLys 773
DB 2338 GGTGTTTTCAGCATATATGACAGTGTATGCTTGAATTCCTTTGTGTTCTCTTAAG 2397
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
DB 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTATTACCGAATCTTTGACAGTGTGCG 2457
QY 794 LysGlnAlaAlaAlaLysAla 800
DB 2458 AAGCAAGCGGCTGCAAAAGCC 2478

```

RESULT 2

ABX93198
ID ABX93198 standard; cDNA; 2824 BP.

AC ABX93198;

DT 29-MAY-2003 (first entry)

DE cDNA encoding corn sugar transport protein #1.

XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.

XX Zea mays.

OS
 XX
 PN US2002178468-A1.

```

XX 28-NOV-2002.
PD
XX
PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08326.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
XX
PS Claim 2; Page 13-15; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other;

```

Alignment Scores:

Pred. No.: 3 81e-298 Length: 2824
 Score: 3517.00 Matches: 684
 Percent Similarity: 95.18% Conservative: 27
 Best Local Similarity: 91.57% Mismatches: 34
 Query Match: 86.31% Indels: 2
 DB: 25 Gaps: 2

US-10-051-909-32 (1-800) x ABX93198 (1-2824)

```

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
DB 238 ATGGGGGGCGCGCTGATGCTGCCATCGCGGCTTATCGGCACTTGCTGACGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
DB 298 GACAAATGCGACAATTGCTGAGCCGCTCTGTACATTAAGAAGAAATTCAACCTGCAGACC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
DB 358 GAGCCTCTGATCGAAGGCTCATCGTCGCAATGTTCTCATTTGGGGCAACAGTCATACCA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
DB 418 ACATCTCCGGGCGCAAGGCTGACTGCGTTGTAGAGGCCCATGCTGTGCTGCGCTGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
DB 478 GTCCTCTACTTCTGTCAGTGGCTGTGATGCTTTGGGGCCCAATTGTATACATCTTGCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
DB 538 CTCGCAAGGCTCATGTGATGGGTTGGTATCGGTTGGCGGTACACACTGTCTCTCTAC 597

```

QY 176 ILeSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAAACTGCACCGCACAGANATTCTTGGGCTGNTNGAACACGTTGCCGCACTTC 657
QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATGGGGGTCAAGGAGGATGTTCTCTCTCTACTGCAATGTTGGATGTCCTCATG 717
QY 214 ProSerProAspTrpArgIleMetLeuGlyValIleuAlaIleProSerLeuPhePhe 233
Db 718 CCCAACCCTGATTGAGGCTCATGCTTGAGTCTGTGCATCCCGTCACCTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTTrpLeuValSerLySGlyArgMet 253
Db 778 GGACTGACTGTCTTCTACTTCCCTGAATCACCAAGGTGGCTGTNAGCAAAGGAGATG 837
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeu 273
Db 838 GCGAGGCGGAAGAGATGNTGCATAAGGCTGCGGGAGAGAGAATGTTCTCANGGAGANG 897
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
Db 898 GCTCTTAGTTGAAGGTTTGGGGGTGCGTAAGATACAGTATTNAGATACATCATTT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db 958 GGAGCTGCCACCGAGCGACCGATGATCTGTAACTGACGGTGATAGAACAATCACA 1017
QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLySGlyProIleMet 333
Db 1018 CTTTATGGGCTGAAGAAGCCAGTCATGATGCTGCACCTTCTAAGGAGCCCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 1078 CTTGGAAGTGCTTCTTCTTGCACTCTGTCATGGAGCATGGTGAACCAAGATGACCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGATCCGATTGTGACATTTTGGTAGTGTCATGAGAATATGCCCTCAAGCTTGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGTATGAGGACACATGTTTCCAACTTTGGAAGTATGTTCAAGTGCACAGATCAG 1257
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1258 CATGCCAAAAATGAGCAGTGGATGAAGAATCTTCAAGGATGACGAGAGATACGA 1317
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGGTGCAGGAGGTGACTATGAGCACAATCTCCATAGCCCATTTGCTGCCAGCAG 1377
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGCGGAAGGGAAGACATTTGTGCACCATGTGCACCGTGGAAGTGTTCAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyGlyAspGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAAGGCMAAGCCTTTTGGGGAGGGGTGAGATGTTGAGCAGCAGCTGATATCGGT 1497
QY 474 GlyGlyTyrGlnLeuAlaTTrpLysTrpSerGlyLysGluGlyGluAsnGlyArgLySGlu 493
Db 1498 GGGGGATGGCAGCTTGCTTGAATGTGTCAGAGAAGAGGTGAGAATGTGTAAGAAAGAA 1557
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgLySer 513
Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCCTGGCTCAAGAGAGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1618 ATGTGTTCACTTCCCGGTGGTGGATGTTCTTGAGGGTATGAGTTTGTACATGCTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLySGlyLeuAlaGluProArgMetSerAsp 553

Db 1678 GCTTAGTAGTACAGTACAGCACTTTCTCAAGGGTCTTGCTGAACCAACGATGTCAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTTrpLysAspLeuPhe 573
Db 1738 GCTGCATGTTTCAACCATCTGAGGTAGCTGCCAAGGTTACAGTTGGAAAGATTGTTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePhe 593
Db 1798 GAACCTGAGGTGAGCGGTGCCCTGTAGTCGGTGTGAATTGAGATCCATTCAACAGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1858 GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGGTGGCA 1917
QY 614 ValIleLeuSerLySPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATTCTTCCAAATTGGTCTCAGCTCGGCATCAGCATCCATTCGATTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACCTACTAATGCTTCTTGCAATGGCTTTGCCATGCTGCTTATGATCTTCCGGA 2037
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGTTTGTGCTGTAAGCACAAATTCCAATCTTGATAGCATCTAGTTATCCTGTT 2097
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAATCTAATGATTGGGTGATACATGCCCATGCTTGTGCTTCCACATCAGTGT 2157
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATGCTACTCTGCTGCTTCTGTTATGGATTGGTCCCATCCCAACATTTTATGTGCA 2217
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
Db 2218 GAGATCTTCCAACCAAGGCTTGTGCGCTGTATTCATTTGTGCTTACATTTCTGG 2277
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGAGATATCATGCTACCTACAGCCTTCTGTGATGCTGAATGCTATTTGGACTGGCG 2337
QY 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheLeuLys 773
Db 2338 GGTGTTTACGATATATGACGTGATGCTGATTTCTTGTGTGCTTCTTAAAG 2397
QY 774 ValProGluThrLySGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAAGGGATGCCCTTGAGGTATTAACGAATCTTTGCAAGTTGGTGGCG 2457
QY 794 LysGlnAlaAlaLysAla 800
Db 2458 AAGCAAGCGGCTGCATAAAGCC 2478
RESULT 3
ABK51965
ID ABK51965 standard; cDNA, 2601 BP.
XX
AC ABK51965;
XX
DT 27-AUG-2002 (first entry)
XX
DE Soybean contig encoding A. thaliana-like sugar transport protein.
XX
KM Soybean; Arabidopsis thaliana-like sugar transport protein;
KM carbohydrate transport; grain filling; annual field crop;
KM plant; gene; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 175..2388

FT	/*tag= a	"Soybean Arabidopsis thaliana-like sugar
FT	/product=	transport protein"
FT		

PN US6383776-B1.

PD 07-MAY-2002.

PF 14-APR-1999; 99US-0291922.

PR 24-APR-1998; 98US-083044P.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

DR WPI; 2002-451386/48.

DR P-PSDB; AAU97204.

PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 3; Column 35-38; 54pp; English.

CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC representing a contig assembled from various soybean cDNA clones
CC encodes an Arabidopsis thaliana-like sugar transport protein.

Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other;

Alignment Scores:

Pred. No.:	2.94e-224	Length:	2601
Score:	2674.00	Matches:	522
Percent Similarity:	82.20%	Conservative:	92
Best Local Similarity:	69.88%	Mismatches:	115
Query Match:	65.62%	Indels:	18
DB:	24	Gaps:	8

US-10-051-909-32 (1-800) x ABK51965 (1-2601)

QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr	75
Db	175	ATGAAAGGTGCCGCTCTTGTTGCTAATGGCCGCTTCCATTGGTAATTTCTCCAGGATGG	234
QY	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95
Db	235	GATTAATGCTACCATCGCCGGGCTAATGGTTACATTAAGAAAGACCCTTGCTTTGGGAACA	294
QY	96	GluProThrValGlnGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	295	-----ACTATGGAAGGCTTGTTGATGGGCATGTCCTGATTGGAGCAACGGTATCACCC	348
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	349	ACATGCTCTGGTCCCTATAAGCGGATTGGCTCGGTGCGGCGAACCCATGATGATATCTCATCT	408
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPserProAsnValTyrValLeuLeu	155
Db	409	GTGCTCTAATTTCTTGGGCTGTTGGTGAATGCTGTGGTCCCAATGTGATGTGTGTGTC	468
QY	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	469	TTGGCGAGGCTACTTGATGGAATTTGGGATTTGGCCCTTGCTGTGACTTGTGTCCCGGTCTAT	528

QY		176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db		529	ATATCTGAACCGCCGCTCTGAAATAAGGGGTCTGTAATACGCTTCTCAGTTCAGT	588
QY		196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Db		589	GGCTCTGAGGAATGTTTTTGTCGTACTGATGTTTTTGGCATGTCAATGATCCCCGC	648
QY		216	ProAspTrpArgIleMetIleuGlyValIleuAlaIleProSerLeuPhePheGlyLeu	235
Db		649	CCTAGCTGAGGCTCATGCTTGGGGTCTGTCTAATTCCTTCTCTGTGATTTGCCATTG	708
QY		236	ThriLepheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255
Db		709	ACCATTTTTCTCTGCCGAGTCTCTCGGTGGCTGTCACAAAGAAGATGCTCGAG	768
QY		256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
Db		769	GCTAAGAAGTGTCTCCAAGATTGCGCGAAGGAGGATGTGTACAGCGCAGATGGCATTG	828
QY		276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
Db		829	CTGGTTGAAGGTCTCGGGAATTGGGGTGATACATCTATCGAAGATACATAATTGGCCCT	888
QY		296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
Db		889	GCTGACGATGTGGCTGATGTGCATGAACAACAGCACAGAAAAGATMAAATTCGATTATAT	948
QY		316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
Db		949	GGATCCCAGACAGCGCTTCTTGTTATCAAACCTGTCACTGCACAGAGTTCTAATTGGC	1008
QY		336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db		1009	-----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCCCTCATG	1056
QY		356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Db		1057	GATCCTCTGTGACACTGTTTGTGATCATGACAGAACCTCCCGACAGACAGGACAAGA	1116
QY		374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db		1117	GGAAGCATGCGAAGCACTCTGTTCCAAATTTTGAAGCATGTTCAGCACCTGTGACCG	1176
QY		394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db		1177	CATGCTAAAAATTGAACAAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG	1236
QY		414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db		1237	TCAATGCAACCCGCTGGGACTCCGATGATTAATTGACAGCTCCTTAATCTACGCCAA	1296
QY		434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu---	452
Db		1297	ACAACAAGCCTTGAA---AAAGACTTACCTCCTCCTCCATGCGACAGTATCCTTGGC	1353
QY		453	SerMetArgArgGlnThrLeuLeuGluGlyGly---GlyAspGlyValSerSerThrAsp	471
Db		1354	AGCATGAGCGCTCACAGTAGTCTCATGCAAGGGTCAAGGTGAGCAAGGTGGTAGTACAGT	1413
QY		472	IleGlyGlyGlyTyrPglLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg	491
Db		1414	ATTGCTGTGGCTGCGCACTGGCATGGAAATGAGACTGATAAA--GGTAGAGATGAAAAA	1470
QY		492	LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg	511
Db		1471	CAACAAGAGGTTTAAAGAGATTATTACATGAGAGGAGGAGTTTCTGCATCTGTCGT	1530
QY		512	GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis	531
Db		1531	GGATCCATTGTATCGATTCCCGGTGAAGC-----GAATTTGTTCAG	1572
QY		532	AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet	551

Db 1573 GCTGCTGCCTTGGTAAGCCACCCGCTCTTACTCCAGAGAGCTTATGTGACACCCA 1632
QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaAlaGlySerArgTProLysAsp 571
Db 1633 GTTGGGCGCTGCATGTGTTCAACCACTGAGACAGCTTCAAGGGGCCAAGTTGGAAGCT 1692
QY 572 LeuPheGluProGlyValArgAlaAlaLeuValGlyValGlyIleGlnIleGln 591
Db 1693 CTCTTGAACCAAGGGTTAAGCATGCTTGGTTGTGGAGTTGGAATACAAATACTTCAG 1752
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnIleGly 611
Db 1753 CAGTTTTCAGGATAAATGGGTTCTATATTACACACCTCAAAATCCTTGAAGAGCGGT 1812
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTTCTTCAAGATAAGCATTTGGCTCAGAGTCGGCATCATTCCTTATCAGT 1872
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCACAACCTCTTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGATGTT 1932
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAGAAAGCAGTGTCTACTTACTACAATCCCCGTCGTGATTGTGTCACTCATATT 1992
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGTCATTTGGAAGCCTGGTAATTTTGGCAATGTGCCCATGACAGCAATCTCAACAGTA 2052
QY 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TGCCTGTGTTATTCTTCTGCTTGTGATGGGTTATGAGCAATTCCAAACATCCTT 2112
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTCCCACTAGGGTGCGCTCGCCTTCGCAATGCTATCTGTGCATTAAGTG 2172
QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGGATTGGAGACATCATCATCACTACTGCTGCTGTGATGCTCGGCTCTTTAGGA 2232
QY 752 LeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPhe 771
Db 2233 CTGGTGTGTATTGCGCAATTACGCAGGTGTTGTTCATCTCGTGATATATTGTTT 2292
QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db 2293 TTGAAGTTTCCAGAAACAAGGGCATGCCCTTGAAATCATCTCTGAATCTTTTCTGTT 2352
QY 792 GlyAlaLysGlnAlaAlaAla 798
Db 2353 GGAGCAAAAGCAGGCTGCTCT 2373

RESULT 4

ABX93201 standard; cDNA; 2601 BP.
XX ABX93201;
AC ABX93201;
XX 29-MAY-2003 (first entry)
DT
XX
DE cDNA encoding soybean sugar transport protein #1.
XX
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KM plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Glycine max.
XX
PN US2002178468-A1.
XX

PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALIE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08329.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein
PS Claim 2; Page 19-20; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other;

Alignment Scores:
Pred. No.: 2.94e-224 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 25 Gaps: 8

US-10-051-909-32 (1-800) x ABX93201 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db 175 ATGAAGTGCCTGCTTGTGCTATTGCGCTTCATTGGTAATTCTCCAAAGATGG 234
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 235 GATAATGCTACCATCGCGGCTTAATGCTTACATTAGAAGAACCTTGCTTGGGAACA 294
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGAAAGGCTTGTGTGGCGCATGTCCCTGATTGGAGCAACGTAATCAC 348
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 349 ACATGCTTGTGCTTATAGCGGATTTGGCTCGGCGGACCATGATGATATCTCATCT 408
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db 409 GTGCTTATTCTTGGGTGTGTGGATGTGTCGCCCAATGTGATGTGTGTGC 468
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 469 TTGGCGAGCTACTTGATGATTTGGGATTTGGCCTTGTGCTGTCCCGGTCTAT 528
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195

```

Db      529 ATATCTGAAACGCGCGCTCTGAAATAAAGGGGCGTGTGATATACGCTTCCAGTTCAGT 588
QY      196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db      589 GGCTCTGAGGAAATGTTTGTCTGTAAGTGTATGCTTTTGGCAATGTCATTGAGTCCCGCG 648
QY      216 ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db      649 CCTAGCTGAGGCTCATGCTTGGGGTCTGTCTATTCCTCTCTGTAATTTCAGATTG 708
QY      236 ThrIlePheTyrLeuProGlySerProArgTyrIleValSerLysGlyArgMetAlaGlu 255
Db      709 ACCATTTTTCTTGTGCGGAGTCTCTCGGTGCTGTGACCAAGGAAGGATGCTCGAG 768
QY      256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyLysLeuSerLeu 275
Db      769 GCTAAGAAGGTGCTCCAAAGATTGCGCGGAGAGGAGATGTGTACGGCAGATGGCATTTG 828
QY      276 LeuLeuGlyGlyLeuGlyValGlyGlyAspThrSerIleGlyLysIleIleGlyPro 295
Db      829 CTGGTTGAAGGTCTCGGAGATTGGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888
QY      296 AlaThrGlyAlaAlaAspAspLeuValThrAspGlyAspLysGlyGlnIleThrLeuTyr 315
Db      889 GCTGACGATGTGGCTGATGGTTCATGAACATGCAACACAGAAAGATAAAATTTCATTTATAT 948
QY      316 GlyProGlyGlyGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db      949 GGATCCCAAGCAGGCTTTCTTGTGTTATCAAAACCTGTCACTGACAGAGTTCTATTGGC 1008
QY      336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db      1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
QY      356 AspProIleValThrLeuPheGlySerValHisGlyAsnMetProGlnAlaGly----- 373
Db      1057 GATCCTCTGTGACACTGTTGGTAGCATTCATGAGAAAGCTCCCGAGACAGAGCAAGA 1116
QY      374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db      1117 GGAAGCATGCGAAGCACTCTGTTCCAAATTTCGAAGCATGTTCAAGCACTGCTGAGCCG 1176
QY      394 HisAlaLysAsnGlyGlnTyrAspGlyGlyAsnLeuHisArgAspAspGlyGlyTyrAla 413
Db      1177 CATGCTAAAAATTGAACATGGGATGAAGAAAGCTTACAAAGGAACGTGAGACTATCATG 1236
QY      414 SerAspGlyAlaGlyLysAspTyrGlyAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db      1237 TCAGATGCAACCCGTGGGAGCTCCGATGATTAATTGCAAGTCCCTTAATCTCAGCCAA 1296
QY      434 AlaThrGlyAlaGlyGlyLysAspIleValHisIleGlyHisArgGlySerAlaLeu--- 452
Db      1297 ACAACAAGCCTTGAA---AAAGACTTACCTCCTCCCTCCCATGGCAGATTCCTTGGC 1353
QY      453 SerMetArgArgGlnThrLeuLeuGlyGlyGly---GlyAspGlyValSerSerThrAsp 471
Db      1354 AGCATGAGGCGTCAACAGTAGTCTCATGCAAGGGGTCAAGGTGAGCAAGGTGGTAGACAGGT 1413
QY      472 IleGlyGlyGlyTyrGlnLeuAlaTyrLysTyrSerGlyLysGlyGlyGlyAsnGlyArg 491
Db      1414 ATGTGTGTGGCTGGCACTGGCATGGAATGAGTGAATAA---GGTAGAGATGAAAAA 1470
QY      492 LysGlyGlyGlyPheLysArgValTyrLeuHisGlnGlyValProGlySerArgArg 511
Db      1471 CAACAAGAGGAGGTTTAAAGATTATTTACATGAGAGGAGGAGTTTCTGCATCTCGTCT 1530
QY      512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGlyGlySerGlyPheValHis 531
Db      1531 GGATCCATTGTATCGATTCGCGGTGAAGG-----GAATTGTCCAG 1572
QY      532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGlyProArgMet 551

```

```

Db      1573 GCTGCTGCCTTGGTAAACCAACCCGCTCTTTACTCCAAAGAGGCTTATGTAGGACACCCA 1632
QY      552 SerAspAlaAlaMetValHisProSerGlyValAlaAlaLysGlySerArgTyrLysAsp 571
Db      1633 GTTGGGCTGCAATGTTTCAACCATCTGAGACAGCTTCAAAAGGGCCAAAGTTGGAAGCT 1692
QY      572 LeuPheGlyProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db      1693 CTCTTGAACCAAGGGGTTAAGCATGATGTTGGTTGGAGTTGGAATACAAATACTTCAG 1752
QY      592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlyGlnAlaGly 611
Db      1753 CAGTTTCAGGAGATTAATGGGTTCTATATTACACCTCAAAATCTTGAAGAGCGCGT 1812
QY      612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db      1813 GTTGAAGTTCTTCTTTCAGATATAGCATGTGGCTCAGAGTCGGCATCATTCCTTATCAGT 1872
QY      632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db      1873 GCTTTCACAACCTTCTGATGCTTCCCTGTATAGCGGTAGCCATGAAAGCTCATGATGTT 1932
QY      652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db      1933 TCAGCAGAGAGGACAGTTGCTACTTACTACATCCCGGTGCTGATGTGTACTCATTTAT 1992
QY      672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db      1993 TTGCTCATTGAAGCGCTGTGTAATTGTGGCAATGTGCGCCATGACGAATCTCAACAGTA 2052
QY      692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db      2053 TGCCTGTGTTATTTCTGCTGCTTGTGTATGGGTTATGGACCAATTCCAACATCCTT 2112
QY      712 CysAlaGlyIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db      2113 TGCTCAGAGATTTTCCCACTAGAGGTGCGTGCCTGTCATTGCTATCTGTGCACTAGTG 2172
QY      732 PheTyrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db      2173 TTCTGATTGGAGCATCATCATCATACATACGCTGCTGTGATGCTCGGCTCTTTAGGA 2232
QY      752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db      2233 CTGTGCTGTATTCGCCATTATACGAGTGTGTTGTTTCATCTCGTGATATTGTGTTT 2292
QY      772 LeuLysValProGlyThrLysGlyMetProLeuGlyValIleThrGlyPhePheAlaVal 791
Db      2293 TTGAAGGTTCCAGAAACAAAGGGCATGCCCCCTTGAAGTCATCTGGAATCTTTTCTGTT 2352
QY      792 GlyAlaLysGlnAlaAla 798
Db      2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 5
ABZ13510
ID      ABZ13510 standard; DNA; 2190 BP.
XX
AC      ABZ13510;
XX
DT      21-JAN-2003 (first entry)
XX
DE      Arabidopsis thaliana stress regulated gene SEQ ID NO 1315.
XX
KW      Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS      Arabidopsis thaliana.
XX
PN      WO200216655-A2.
XX
PD      28-FEB-2002.
XX
PF      24-AUG-2001; 2001WO-US26685.

```


XX 24-AUG-2000 2000US-227866P.
PR 26-JAN-2001 2001US-264647P.
PR 22-JUN-2001 2001US-300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT
XX
PS Claim 144; SEQ ID NO 1315; 577bp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 2190 BP; 529 A; 439 C; 572 G; 650 T; 0 other;

Alignment Scores:
Pred. No.: 1.79e-216 Length: 2190
Score: 2584.00 Matches: 512
Percent Similarity: 80.78% Conservative: 89
Best Local Similarity: 68.82% Mismatches: 123
Query Match: 63.41% Indels: 20
DB: 24 Gaps: 8

US-10-051-909-32 (1-800) x AB213510 (1-2190)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
DB 1 ATGAGTGGAGCTGTGCTTGCTTATTGCTGCTGCTGCTGGCAACTTGTACAAAGATGG 60
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleValysGluPheGlnLeuGlnAsn 95
DB 61 GATAACGCAACTATTGCAGAGCTGTGTGTACATMAAAGAGATTAAATTGGAGACT 120
QY 96 GluProThrValGluGlyLeuIleValSerMetSerIleGlyAlaThrIleValThr 115
DB 121 AATCCATCAGTGAAGGCTTAATTGGCGATGTCACTAATTGGTCTACTCTGATTACA 180
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
DB 181 ACATGCTCTGGAGGGGTAGCTGATTGGCTTGTCGCCCTCCATGCTAATATTGTCCTCA 240
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
DB 241 ATTCTCTACTTGTGTGCTCTAGTAATGCTATGCTCCGAATGTTAAGTGTGCTC 300
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
DB 301 TTAGGAAGGTTGTAGATGATTGGGGTGTGCTTGTGGTCAACACTTGTCTTATTTAT 360
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
DB 361 ATATCTGAGACTGCACCACTGAGATTAGGGGACTGTGAATAACGCTACCGCAGTTTCACT 420

QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
DB 421 GGCTCTGGAGGAGATGTTCTTATCTTACTGTATGCTTTTCGAATATGCTGTGATGCCATCA 480
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
DB 481 CCTAGCTGAGATGATGCTTGTGTGCTTTCATCCCTTCCCTTGTCTTTTCTTCTC 540
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIlyArgMetAlaGlu 255
DB 541 ACGGCTCTCTCTTCCGCCAGAGTCCGCCAAGGTGCTGTGAGCAAGGTCGAATGCTTGA 600
QY 256 AlaValysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
DB 601 GCAAAGCGGTTCTTCAGAGACTGCCGTGTCGCCGAAGATGTGTGTTGAGATGCTTTG 660
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
DB 661 TTGTTGAGGGTCTTGAATTGGAGTGAACCAACATAGAGAAATATATATGTTGCTCC 720
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
DB 721 GCGGATGAAGTTACTGATGATCATGATATAGCTGTGATTAAGATCAAAATTATATAT 780
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
DB 781 GGTGCAGAGAAGAGGCTGAGTTGGTGTGCTGAGCCCACTCAAGA-----GGA 828
QY 336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
DB 829 AGCACTATGAGTGTGTTGCTCTGCCATGGAAGTACAATGAGCAGAGGCAAGCTCATTTG 888
QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
DB 889 ATTGATCCTTGTTCACACTGTTGGAGCGTTTCACGAGAAGATGCCGACACT---GGA 945
QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
DB 946 AGCATGAGAGATGCTTGTCCACATTTGGAGTATGTTCACTGTGGAAGGAATCAA 1005
QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
DB 1006 CCAAGACATGAAGATTGGAGTGAAGAATCTTGTGAGAAGGTGAGATTATCCATCC 1065
QY 415 AspGlyAlaGlyLysAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
DB 1066 GAC---CATGAGATGATTTCTGAAGATGATCTTCACTCCGTGATCTCACGTCACAAAG 1122
QY 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
DB 1123 ACAAGCATGAG---AAAGACATGCCCTCACACTGCTCATGGAACCTTTTCTACCTTACA 1179
QY 455 ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
DB 1180 CATGGAAGTCAAGTGCAGGGAGCTCAAGGGGAAGAGCGGTAGTATGGGAGATTGAGGT 1239
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlyGly 494
DB 1240 GGATGGCAAGTGCATGGAATGACGGAAGAAGATGAATCGGACAGAAAGAA--- 1296
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
DB 1297 -----GAGGTTTCCAGAGATCTCGACGTGGCTCAATT 1329
QY 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534
DB 1330 GTTTCATGCTGCTGTGTGATGAACCGGTGAG---GCAGATTGTTGACAAACGCTGTGCT 1386
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
DB 1387 TTGTTAGCCAACCAAGCTCTTATTTCCAAGACCTTCTCAAGAACATACAAATTGCTCT 1446
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574

Db 1447 GCTATGGTACATCCATCCGAA---ACAACTAAAGGGTCAATTGGCATGATCTTCATGAT 1503
QY 575 ProGlyValArgAlaLeuLeuValGlyValGlyLeuGlnLeuGlnPheAla 594
Db 1504 CCTGAGTCAAGCGTGCATTAGTCGTAGAGGTGGACTTCAATACTCAGCAGTTCTCA 1563
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAlaVal 614
Db 1564 GGCATCAACGAGATTCTTACTACACACCGCAATCTTGAGCAGCGGGTGTGGGATC 1623
QY 615 IleuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
Db 1624 CTACTATCGAACATGGGAGTTAGTTCTTCTCAGCATCTTACTTATAGTGCATTGACA 1683
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
Db 1684 ACCTTTGTGATGTATCCTGCATAGCTGTGCAATGAGGCTCATGATCTTCTGGTGA 1743
QY 655 ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
Db 1744 AGGACCTTGCTCTCAACCATGATCCATCTGATAGCATCTTATGTTTATGATATC 1803
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
Db 1804 TCAATCTTGTTCACATGAAACAGCATTTGTGCACGCGGTCTTATCAACCGTAAGCGTGTG 1863
QY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
Db 1864 CTCTACTTCTGCTTCTTCTGATGGGTTTGGTCTCCTGCTCAAAACATCCTCTGTCAGAG 1923
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
Db 1924 ATTTTCCAACTCGAGTCGCGGAAATCTGCATCGCCATCTGCGCACTCACCTTGTGATC 1983
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
Db 1984 TGTGACATTAATCGTCACTTACAGTCTCCCGTGTCTCAAAATCCATTGGACTAGCTGCT 2043
QY 755 ValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLysVal 774
Db 2044 GTGTTTGGAAATGATCGCAATCGTATGTTGCATTTCATGGGTCTTGTGTTCAATTAAGTC 2103
QY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
Db 2104 CCGGAAACTAAAGCATGCGACTTGAAGTCATCACAGAGTTCTTCTGTTGAGCTAGA 2163
QY 795 GlnAlaAlaAla 798
Db 2164 CAAGCTGAAGCT 2175
RESULT 6
ABZ14449
ID ABZ14449 standard; DNA; 2205 BP.
XX AC ABZ14449;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.
XX KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US26685.
XX PR 24-AUG-2000; 2000US-227866P.
XX PR 26-JAN-2001; 2001US-264647P.

PR 22-JUN-2001; 2001US-300111P.
XX (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT
XX
PS Claim 144; SEQ ID NO 2254; 577bp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 other;
Alignment Scores:
Pred. No.: 2.29e-189 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
DB: 24 Gaps: 14
US-10-051-909-32 (1-800) x ABZ14449 (1-2205)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 1 ATGAAGGAGGCGACTCTCGTTGCTCTCCGCCCAATCGCAATTCTTCAAGATGG 60
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 61 GACAAATGCCACCATGCTGAGAGCTATGTTATATCAACAAGACTTGAACTA----- 114
QY 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CCAACCTCTGTTCAAGGTTCTGTGCTGATGTCATTGATCGTGCAACGCTCATC 171
QY 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuLeuSer 134
Db 172 ACGACTTGCTCAGGACCATATCTGATTGGCTCGGAGAGAGCCCATGCTCATTTATCA 231
QY 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
Db 232 TCAGTTATGATTTCGTCTGCGGTTTGATTAATGTGTGCTCCCAATGTCTATGTTCTG 291
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 292 TGCCTTGTAGGCTTCTTAATGAGTTTGTTGTCGCGGCTCGCGTTACACTTGTCCCTGT 351
QY 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 352 TACATTTCTGAACCGCTCTCCGAGATCAGAGACAGTTAAATACTCTCCCTCAGTTT 411
QY 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 412 CTGGCTCTGTGAATGTTTGTGTCATAGTATGTTTCACTATGTCCTGAGTGAC 471

OY		215	SerProaspTTrpargIlemetLeuGlValLeuAlaIleProSerLeuphepheGly	234
Db		472	TCCCTAGCTGGAGAGCCATGTCGGTGTCCTCGATCCCTTCTTTATTGTGTT	531
OY		235	LeuThrIlePheTyrlLeuproGluSerProargTrpleuValserLySglYargmeAla	254
Db		532	CTCACGGTGTATTATTGCCCGAGTCTCCGTGTGCTGGTTAGTAAGAAGAAATGCAC	591
OY		255	GIUAlalysLysValLeuGlnLysLeuArgGlyLYsaAspaPvalSerGlyluleUser	274
Db		592	GAGGCTAAGCGAGTCTTCAACA GTTATGTGGCAGAGAAAGATGTTACC GATGAGATGGCT	651
OY		275	LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerileGluGluTyrlleIegly	294
Db		652	TTAATAAGTTGAAGACTAGATATAGAGAGAAAACAATGGAAGATCTCTTAGTA ACT	711
OY		295	ProAlaThrGUAlaAlaAspaSbleuValThraspGlyAspLysGlu--GlnIleThr	313
Db		712	TTGGAGGATCATGAAGGTGATGATACACTTGA AACCGTTGATGAGATGACAAATGCGG	771
OY		314	LeuTyrglyProGluGluGlyInserTrpIleAlaArgProSerLySGlyProIleMet	333
Db		772	CTTATGGAACCCAGAGAAATCAATCGTACCTTGCTAGACTGTCCCAGAACAA-----	825
OY		334	LeuGlySerValLeuSerLeuAlaSerArghisGlySerMetValasngInserValPro	353
Db		826	--AATAGCTCACTTGGGCTACGCTCTGCCACAGGAAAGCTTACCAAACCAAGCATGATC	882
OY		354	LeuMetaspProIleValThreupheGlySerValHisGluasnMetProGlnAlagly	373
Db		883	CTTAAGATCCGCTCGTCAATCTTTTGGCAGCTCTCCACGAAAGATGCCAGAAAGCAGC	942
OY		374	GlySerMetArgSerThrLeupheProasnphelGlySerMetPheserValThraspin	393
Db		943	GGAACAACCTGGAGTGGGATTTCCCTCATTTCCGAAAGCATGTTCACTACTGCGCAT	1002
OY		394	----HisAlalysAsnGluIntPrasp-----GluGluAsnLeuHisArgAsp	408
Db		1003	GCGCCTCACGGTAACCGGCTCATTTGGGAAAGACATAGAGAGCCATTACACAAAGAC	1062
OY		409	AspGluGluTyrlaserAspGlyAlaglyGlyAspTyrcIuaSp-----AsnLeu	425
Db		1063	AATGATGACTATGCGACTGATGATGTCGGGTGATGATGATGATCGGACAACGATTTG	1122
OY		426	HISerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis	445
Db		1123	CGTAGCCCCCTTAATGTCCGCCAGACACACAAGCATGGAC--AAGGATATGATCCCATAT	1179
OY		446	GlyHisArgGlySexAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyAsp	465
Db		1180	CCTACAAGTGAAGCATTTAAGCATGAGACGACACACAGTACGCTTATGCAA--GGCAAC	1236
OY		466	GlyValSerSerThrAspIleGlyGlyGlyTyTrpGlnLeuAlaTrplySTPserGlyLys	485
Db		1237	GCGCAAAAGTAGCATGGGAATGTGTGTTGGCAATATGGGATATAGATACGAAAAAGCAT	1296
OY		486	GluGlyGluAsnGlyArgGlyGlyGlyPhelyArgValTyrlEuHisGlnGlyGly	505
Db		1297	GAA-----TACAAAGGTATTATCTTAAGAAAGATGGA	1329
OY		506	ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValpheGlu	525
Db		1330	GCTGAA--TCTCGCCGTGGCTCGATCATCTCTATGCCGAGATCCCGAT-----GGT	1380
OY		526	GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaleuPheSerLySgly	545
Db		1381	GGAAGCAGCTACATTCACGCTTCTGCCCTGTAAAGCAAGATCTGTTCTTGCTCTAAATCA	1440
OY		546	LeuAlaGluProArgMetSerAspAlaAlametValHisProSerGluValAlaAlaLys	565
Db		1441	-----GTTATGGAATCCGCCATGTGTTCCCGGAGAAATGTGCTGCTCT	1485

Accession	Gene	Protein	Length	Source
QY	566	GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuValGlyVal	585	Arabidopsis thaliana
Db	1486	GGACCACCTCTGGTCTGCTCTTCTTGAACTGGTGTAAAGCGTGCCTTGGTGTGTC	1545	Arabidopsis thaliana
QY	586	GlyTLeuGlnIleLeuGlnIlePheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln	605	Arabidopsis thaliana
Db	1546	GGCATTCAATATACGACGACGATTTTCAGGTATCAATGAGTTCTTACTACCTCTCAG	1605	Arabidopsis thaliana
QY	606	IleLeuGlnIleGlnIleValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer	625	Arabidopsis thaliana
Db	1606	ATTCTCGAAGCGGCTGGCGTAGATATCTTCTTCGAGCCTCGACTAAGTTCATCTCT	1665	Arabidopsis thaliana
QY	626	AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla	645	Arabidopsis thaliana
Db	1666	GGCTCATCTCCTCATCAGCGGTTTAAACAATATCTCATGCTCCAGCCATGTGCTGCC	1725	Arabidopsis thaliana
QY	646	MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeu	665	Arabidopsis thaliana
Db	1726	ATGAGACTCATGATGATATCCGGAAGAGGTCAATTACTTCTTGACAATCCCACTTCTC	1785	Arabidopsis thaliana
QY	666	IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrIleAlaHis	685	Arabidopsis thaliana
Db	1786	ATTGTCTACCTTGTCTCTTGTCTCATCAGCGAGCTCATCCACATCAGCAAAAGTCGTAAC	1845	Arabidopsis thaliana
QY	686	AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly	705	Arabidopsis thaliana
Db	1846	GCAGCACTCTCCACAGGTTGTGTGCTGCTCTTCTGCTTCTTCTGATGGGTAACTGCT	1905	Arabidopsis thaliana
QY	706	ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle	725	Arabidopsis thaliana
Db	1906	CCCATTCCAAACATCCTCTGTCTGTAATCTTCCCAACAAGAGTCCGTGTCTCTGCATC	1965	Arabidopsis thaliana
QY	726	AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal	745	Arabidopsis thaliana
Db	1966	GCCATATGTGCTATGCTCTTTTGGATTGGAGACATTATTGTACAGTACTTCTCCCGTT	2025	Arabidopsis thaliana
QY	746	MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValCysLeuIle	765	Arabidopsis thaliana
Db	2026	CTCCTCAGCTCGATCGAGCTAGTGTGTGTTTTCAGCAATTACGCTGCGGTTTGCCTTATC	2085	Arabidopsis thaliana
QY	766	SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle	785	Arabidopsis thaliana
Db	2086	TCATGATCTTCGTTTACATGAAGAATCCCGAGACTAAAGGCATGCTTTGGAAGTTATC	2145	Arabidopsis thaliana
QY	786	ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla	800	Arabidopsis thaliana
Db	2146	ACAGACTACTTGGCTTTGGAGCT--CAAGCTCAAGCTTCTGCT	2187	Arabidopsis thaliana
RESULT 7				
ID	ABK51966	standard; cDNA; 1692 BP.		
XX	ABK51966;			
DT	27-AUG-2002	(first entry)		
XX				
DE	Soybean cDNA clone ss1.pk0022.fl encoding sugar transport protein.			
KM	Soybean; Arabidopsis thaliana-like sugar transport protein;			
KM	carbohydrate transport; grain filling; annual field crop;			
KM	plant; clone ss1.pk0022.fl; gene; ss.			
XX				
OS	Glycine max.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	9..1469		
FT		/*tag= a		
FT		/partial		
FT		/product= "Portion of a soybean Arabidopsis		
FT		thaliana-like sugar transport protein"		
FT		/note= "This sequence lacks a start codon"		
XX				
PN	US6383776-B1.			

XX 07-MAY-2002.
XX 14-APR-1999; 99US-0291922.
XX 24-APR-1998; 98US-083044P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2002-451386/48.
XX P-PSDB; AAU97205.
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX Claim 3; Column 41-44; 54pp; English.
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence encodes
CC a portion of a soybean Arabidopsis thaliana-like sugar transport
XX protein.
SQ Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 other;

Alignment Scores:
Pred. No.: 1.63e-136 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 24 Gaps: 8

US-10-051-909-32 (1-800) x ABK51966 (1-1692)

QY 308 AsplysGluGlnIleThrLeuTyrglyProgluGluGlnSerTrpIleAlaArgPro 327
Db 21 GAAAAAGATCAATTAAGTTGATGACCAAGAACAGCCAGCTCGGTGCTAGACT 80
QY 328 SerlySGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db 81 GTTGCTGACCAAAATTCGTGGC-----CTTGATCTAGGAAAGGAGCATG 128
QY 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367
Db 129 GCAAAATCCAAGC--AGTCTAGTGACCCCTCTAGTGACCCCTTTGGTAGTACATGAG 185
QY 368 AsnMetProGlnAlaGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCAGAAACAGAG-----AGCACCCCTTTTCCACACTTTGGAGTATG 233
QY 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
Db 234 TTCAGTGTGGGGGAATCAGCCCAAGAAATGAAGATTGGGATGAGGAAAGCCTAGCCAGA 293
QY 408 AspAspGluGluTyralaSerAspGlyAlaGlyIleAspTyrgluAspAsnLeuHisSer 427
Db 294 GAGGTGATGATTATGCTCTGAT-----GCTGTGATTCGATGACAAATTTGCAGAGT 347
QY 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447
Db 348 CCATGTACTCAGTCACAACAGAGTCTGAT--AAGGACATACCTCCTCATGCCCAT 404
QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu--GlyGluGlyIleAspGly 466

Db 405 AGTAACCTTGCA--AGCATGAGGCAAGGTAGTCTTTTACATGGAATTCAGAGAACCC 461
QY 467 ValSerSerThrAspIleGlyGlyGlyTyrcpGlnLeuAlaTrpIleSerGlyLysGlu 486
Db 462 ACTGTAGTACTGGGATTTGGTGTGGTGGCAGCTAGCATGGAATGTCTGAAAGAGAG 521
QY 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrluHisGlnGluGlyVal 506
Db 522 GGCCAGATGGAAGAAGAGGTGGCTTCAGAGAATATATTACACCAAGATGTGTGT 581
QY 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyIleAspValPheGluGly 526
Db 582 TCTGATCTAGACGTGGGTCTGTGTTCACTCCCT--GGCGTGAATTCACCACTGAC 638
QY 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
Db 639 AGTGAGTTGTACAGGCTGCTGCTGTGATGACGCTGCCCTTATATAGAGACCTT 698
QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaLysGly 566
Db 699 ATGCGTCAACGGCCAGTTGGACCAAGCTATGATTCATCCCTCGAAACAAATTGCCAAAAGGG 758
QY 567 SerArgTrpLysAspLeuPheGluProGlyValAlaArgAlaLeuLeuValGlyValGly 586
Db 759 CCAAGTTGAGTATCTTTTGAACCTGGGGTGAAGCATGATGATTGTGGGGGTGGGA 818
QY 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrrTrpProGlnIle 606
Db 819 ATGCAAAATCTTCAGAGTCTCTGTGATAAATGGGGTCTCTACTATACGCTCAAAATT 878
QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTTGAGCAGGAGGTGTGTTATCTTCTTCAAGCCTTAGGCCCTGTGTTCTTACTTCTTCA 938
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTCTTATTAGTCGGGTGACACACCTGTGATGCTTCTGTATAGCCATTGCCATG 998
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCATGATATTCAGGACGACAGAGGACCTTGCTCAGTACAAATCCCCGCTTAATA 1058
QY 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTCTCATATTAGTCCCTGGAGGTCTTGATTTGGATTCACCTGCAAAATGCA 1118
QY 687 LeuLeuSerThrValSerValIleValTyrrPheCysCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATTTAGTATTGTTCTATTTCTGTTCTTGTGATGGGATTGGACCA 1178
QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCTATATATCTTTGTCAGAGATCTTCCCACTCGAGTTGCTGTCATTTGCTT 1238
QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrrSerLeuProValMet 746
Db 1239 ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTCACCTACACACTCCAGTTATG 1298
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyralaValCysLeuIleSer 766
Db 1299 CTCATTTCTGTAGCCCTCGCTGTGTTTGGTATTATGCTGTGCTGTGCTCATAGCA 1358
QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTGTCTTTTGAAGTTCAGAAACCAAGGCGCATGCACTGGAAGTGCATCATT 1418
QY 787 GluPhePheAlaValGlyAlaLysGln 795
Db 1419 GAGTCTTCTCTGTGCGAGCAAAACAG 1445
RESULT 8
ABX93202

Db 1179 ATTCTAATATACCTTTGTCAGAGATCTTCCCTGAGTTCGTGCTCTGCAATTGCT 1238
QY 727 ILeCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746
Db 1239 ATTTGTGCCCTTACCTTTTGATCTGTGATATCATGTGCACCTACACACTCCCAATTATG 1298
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSer 766
Db 1299 CTCATTTCTGAGGCGCTGCTGTTTGTGATTTATGCTGTGCTGCTTCATAGCA 1358
QY 767 PheValPheValPheLeuValProGluThrIleGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTGTCTTTTGAAGTTCAGAAACCAAGGCGATGCCACTGGAAGTATCATT 1418
QY 787 GluPhePheAlaValAlaAlaGln 795
Db 1419 GAGTTCTTCTGTGCGAGCAAAACAG 1445
RESULT 9
ABK51968
ID ABK51968 standard; cDNA; 1487 BP.
XX
AC ABK51968;
XX
DT 27-AUG-2002 (first entry)
XX
DE Wheat contig encoding Arabidopsis thaliana-like sugar transport protein.
XX
KW Wheat; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant; gene; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 3..1040
FT /*tag= a
FT /partial
FT /product= "Portion of a wheat Arabidopsis
FT thaliana-like sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-0291922.
XX
PR 24-APR-1998; 98US-083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-451386/48.
DR P-PSDB; AAU97207.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 3; Column 47-50; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC representing a contig assembled from various wheat cDNA clones encodes

CC a portion of an Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other;
Alignment Scores:
Pred. No.: 5.63e-115 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 24 Gaps: 2
US-10-051-909-32 (1-800) x ABK51968 (1-1487)
QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481
Db 12 GAGGTGGGAGGAGCATGACAGCAGCACTGTATGTGGGGGTGGCAACTCGCATGAAA 71
QY 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGAGCGACAAAGCGAGCATGCGAAGAGAGAGAGGCTTCAAAAGAACTACTTG 131
QY 502 HisGlnGluGlyValProGlySerArgArgLysIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGGTGGCCGCACTCAAGAGGGGCTCTGTGTTCACTTCCGTGGGGGT 191
QY 522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGCGAGTGGGTTTATCATGCTGCTGTTGTAAGCCACTCGGCT 251
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGATCTTATGGAAGAGCGTATGGCGCGCTCCAGCCATGATCATCCA 311
QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGAGGCGAGCTCCCAAGGTTCAATCTGGAAGATCTGTTGAACCTGCTGAGGCGT 371
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTTGTCGGGTGTGGAATTCAGATGCTTCAGCAATTGCTGGAATAAATGAGATT 431
QY 600 LeuTyrTrpThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTATACTCTCAAAATTCGAGCAAGCTGTGCTGCTCTTCTTCCAAATCTT 491
QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCCTCAGTTCAGCATCAAGCATCCATCTGATCACTCTCCACCACCTTACTGCTC 551
QY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATTTGGTGTAGCATGAGACTTATGATATATCTGGAAGAAGATTCTGCTACTG 611
QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAATTCCTCATCTTGATAGCATCCCTAATGTGTTGGGTGTGTCATATATCAAC 671
QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGTGGCCCAAGCTGTGCTCTCCACAGTTAGCGTCATTTGTCTACTGCTGC 731
QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTATGGGCTTTGGCCCGATCCCAACATTTATGTGCAGAGATTTTCCCAACAGA 791
QY 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db 792 GTCCGTGTGTGATCGCTATTGCGCCCTCAATCTGATTTGTGACATTAATGTT 851
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 852 ACCTACAGCCTGCTGTGATGCTGAATGCTATGTGCTAGCGGGTGTCTTGATATATAT 911

QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 912 GCAGTCGTTGCTGCATTCGCTTGTGTTCTGCTAAGGTCCAGAGACAAAGGCC 971
QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAla 798
Db 972 ATGCCCTCGAGGTCAACGAGTCTTTGCGGTTGGGCGAAGCAAGCGCAGGCC 1028
RESULT 10
ABX93204
ID ABX93204 standard; cDNA; 1487 BP.
XX AC ABX93204;
XX DT 29-MAY-2003 (first entry)
XX DE cDNA encoding wheat sugar transport protein #2.
XX KM Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KM plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX OS Triticum aestivum.
XX PN US2002178468-A1.
XX PD 28-NOV-2002;
XX PF 17-JAN-2002; 2002US-0051902.
XX PR 24-APR-1998; 98US-083044P.
XX PR 14-APR-1999; 99US-0291922.
XX PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX DR WPI, 2003-340957/32.
XX DR P-PSDB; ABU08332.
XX PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein
XX PS Claim 2; Page 25-26; 56pp; English.
XX CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other;
Alignment Scores:
Pred. No.: 5.63e-115 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 25 Gaps: 2

US-10-051-909-32 (1-800) x ABX93204 (1-1487)
QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyIleThrGlnLeuAlaTrpLys 481
Db 12 GAGGTTGGGAGGAGCAGTACAGACACAGTGTATGGTGGGGGTGCAACTGCATGAAA 71
QY 482 TrpSerGluLysGluGlyLysGluGlyLysGluGlyLysGluGlyLysGluGlyLys 501
Db 72 TGGTCGAGCGACAGCGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 131
QY 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 521
Db 132 CACCAAGAGGGGGTGGCCGACTCAAGAGGGGCTGTGTTTCACTTCTGTTGGGGGT 191
QY 522 AspValPheGlu--GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGCGAGTGGGTTTATACATGCTGCTGCTTGTGTAAGCACTCGGCT 251
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet--SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGATCTTATGAGAGCGATGCGCGCGGTCAGCCATGATTCATCCA 311
QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGAGAGCAGCTCCCAAGGTTCAATCTGAAAGATCTGTTGAACCTGCTGAGGCGT 371
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTTGCTCGGTTGGAATTCAGATGCTTCAGCAGATTGCTGGAATAATGAGATT 431
QY 600 LeuTyrTyrThrProGlnIleLeuGlnGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTATACTCTCAATTCCTGAGAGCAGCTGCTGCTGCTGCTTCTTCTTCAATCTT 491
QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuMetLeu 639
Db 492 GGCCTCAGTTACGATCAGCATCCATCTTGATCAGTCTCTCACCACCTTACTCATGCTC 551
QY 640 ProCysIleGlyPheAlaMetLeuMetLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATTTGGTGTAGCCATGAGACTTATGATATATCTGGAAGAAGGTTCTGCTACTG 611
QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACATTTCCATCTTGATAGCATCCCTAATGTTTGGGTGTGTCATGTATCAAC 671
QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGGTGCCCCACGCTGTGCTCTCCACAGTTAGCGTCATTGTCTACTTCTGCTGC 731
QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTATGGGCTTTGGCCCCGATCCCCCAACATTTCTATGTGACAGAGATTTTCCACACAGA 791
QY 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db 792 GTCCGTGTGTGTCATCGCTATTGCGCCCTCACAATTGGAATTGTGACATTAATTGTT 851
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 852 ACCTACAGCCTGCGCTGATGCTGATGCTATGCTTAAGGAGGAGGAGGAGGAGGAGGAG 911
QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 912 GCAGTCGTTGCTGCATTCGCTTGTGTTCTGCTAAGGTCCAGAGACAAAGGCC 971
QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAla 798
Db 972 ATGCCCTCGAGGTCAACGAGTCTTTGCGGTTGGGCGAAGCAAGCGCAGGCC 1028
RESULT 11
ABK51969
ID ABK51969 standard; cDNA; 1009 BP.

XX AC ABK51969;
XX 27-AUG-2002 (first entry)
DE Wheat cDNA clone wreln.pk0006.b4 encoding sugar transport protein.
XX
KW wheat; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant; clone wreln.pk0006.b4; gene; ss.
XX
OS Triticum aestivum.
XX
FH Key location/Qualifiers
FT CDS 2..688
FT /*tag= a
FT /partial
FT /product= "Portion of a wheat Arabidopsis
FT thaliana-like sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
PN US6383776-B1.
XX 07-MAY-2002.
XX
PF 14-APR-1999; 99US-0291922.
XX
PR 24-APR-1998; 98US-083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-451386/48.
DR P-PSDB; AAU97208.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 3; Column 51-54; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence encodes
CC a portion of a wheat Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 1009 BP; 237 A; 233 C; 234 G; 305 T; 0 other;

Alignment Scores:
Pred. No.: 4.91e-73 length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
DB: 24 Gaps: 0

US-10-051-909-32 (1-800) x ABK51969 (1-1009)

QY 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 2 GAACCTGAGTGAAGCATGCACCTGTCCTGGCATAGATTCAGATCCTGCAGAGTTT 61
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
Db 62 GCGGTAATCAATGAGTCTCTACTACACACCTCAGATACTTGAGCAAGAGGTGCGGG 121

QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 122 GTTCTTCTATCAACATTGACTAAGCTCTTCTCAGCATCTATCTTATAGTCCTTG 181
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSergly 653
Db 182 ACAACCTTGCTGATGCTTCCAGCATGTGCATGCCATGAGACTCATGATGATGTACAGA 241
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 242 AGAAGTTTCTTCTCTTCAACAATCCCTGTCTGATAGTAGCGCTAGCTGTGTT 301
QY 674 ValSerAsnLeuIleAspLeuGlyThrIleAlaHisAlaLeuLeuSerThrValSerVal 693
Db 302 TTAGTGAATGTTCTGATGTCCGAACCATGTGCACCGCTGCGCTCAACGATCAGCGTC 361
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 362 ATCGTCTATTCTGCTTCTTCTGTCATGGGGTTGGGCTATCCCAATATTCTTGCGCG 421
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 422 GAGATTTTCCCACTCTGTCCGTGCATGTGCATAGCCATCTGCGGCTAACCTTCTGG 481
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 482 ATCGGCGACATCATGTCATACACTCTCCCGTGATGCTCATGCCATGTGCTCGCT 541
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 542 GGAGTCTTCGGCATATATGCCATCGTTGTGTACTAGCCTTGTATTCGTCTACATGAAG 601
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 602 GTCCCTGAGACAAAGGCGATGCCCTGGAGGTATCACCGAGTTCTTCTGTGCGGGCA 661
QY 794 LysGln 795
Db 662 AAGCAG 667

RESULT 12
ABX93205
ID ABX93205 standard; cDNA; 1009 BP.
XX
AC ABX93205;
XX
DT 29-MAY-2003 (first entry)
XX
DE cDNA encoding wheat sugar transport protein #3.
XX
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Triticum aestivum.
XX
PN US2002178468-A1.
PN
PD 28-NOV-2002.
PD
PF 17-JAN-2002; 2002US-0051902.
PF
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALTE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.

DR P-PSDB; ABU08333.
XX Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
XX
PS Claim 2; Page 27; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 1009 BP; 237 A; 233 C; 234 G; 305 T; 0 other;

Alignment Scores:
Pred. No.: 4.91e-73 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
DB: 25 Gaps: 0

US-10-051-909-32 (1-800) x ABX93205 (1-1009)

QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyLeuGlnLeuGlnPhe 593
Db 2 GAACCTGAGTGAAGCATGCACTGTCGTGGCATAGATTAAGATCCTGCAGCGTTT 61
QY 594 AlaGlyIleAsnGlyValLeuYrThrProGlnIleLeuGlnAlaGlyValAla 613
Db 62 GCGGGTATCAATGAGTCTCTACTACACACCTCAATATTGAGCAAGAGGTGCGGG 121
QY 614 ValIleuSerLySpheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 122 GTTCTTCTATCAACATTGGACTAAGCTCTTCTCAGCATCTATTCTTATTAGTGGTTG 181
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 182 ACAACCTTGCTGATGCTTCCCAAGCATGGCATCGCCATGAGACTCATGATATGTCAGGA 241
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 242 AGAAGTTTCTTCTCTCTTCAACATCCCTGTCTTATAGTAGCGTAGCTGTCTGTT 301
QY 674 ValSerAsnLeuIleAspLeuGlyThrIleAlaHisAlaLeuLeuSerThrValSerVal 693
Db 302 TTAAGTGAATGTTCTGTGATGTCGGAACCATGGTGCAACGCTGCCCTCTCAACGATCAGCGTTC 361
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 362 ATGCTCTATTCTGCTTCTTCTGATGGGGTTTG33CCTATCCCAATATTTCTCTGCGCG 421
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 422 GAGATTTTCCCACTCTGTCCGCGCATGTGCATAGCCATCTGCGCGCTAACCTTCTTG 481
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 482 ATCGGCGACATCATCGTACATACCTCTCCCGTATGTCATATGCCATTTGTCGCT 541
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuIys 773
Db 542 GAGGTCTTCGGCATATATGCCATCTGTTGTGTACTAGCCTTTGTATTCTGTACATGAAG 601

QY 774 ValProGluThrLySGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 602 GTCCCTGAGACAAGGCGCATGCCCTGAGGTCATCACCGAGTTCTTCTGTGCGGCA 661
QY 794 LysGln 795
Db 662 AAGCAG 667

RESULT 13
ABL41880
ID ABL41880 standard; DNA; 1518 BP.
XX
AC ABL41880;
XX
DT 11-JUN-2002 (first entry)
XX
DE Nucleotide sequence of a linear polyol transporter.
XX
KW Linear polyol; mannitol; polyol transporter; carbon source; plant;
KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;
KW inositol; ribitol; xylitol; ss.
XX
OS Bacillus subtilis.
XX
PN WO200204647-A1.
XX
PD 17-JAN-2002.
XX
PF 22-JUN-2001; 2001WO-FR01979.
XX
PR 11-JUL-2000; 2000FR-0009032.
XX
PA (CNRS) CENT NAT RECH SCT.
XX
PI Lemoine RRP, Noiraud NEJ;
XX
DR WPI; 2002-154933/20.

XX
XX New polyol transporter protein from plants, for selecting transformed
PT cells and for imparting pathogen and salt-stress resistance to plants
PT
PT
XX
XX Disclosure; Page 46; 66pp; French.
PS
XX
XX The present sequence encodes a polypeptide which is a transporter of
CC a linear polyol. DNA encoding linear polyols is used to produce a
CC selection system for transformed cells, based on the use of polyols
CC as the only carbon source. It is also used to produce transgenic plants
CC with increased resistance to pathogens and salt stress. The use of
CC a linear polyol for selection eliminates the need for toxic selection
CC reagents such as antibiotics. Also, the linear polyol is not essential
CC for the plant, once selection has been made. The linear polyol has a
CC main chain of 5-8, preferably 6, carbon atoms and is selected from
CC sorbitol, dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol.
XX
SQ Sequence 1518 BP; 392 A; 309 C; 389 G; 428 T; 0 other;

Alignment Scores:
Pred. No.: 1.36e-41 Length: 1518
Score: 587.50 Matches: 168
Percent Similarity: 36.64% Conservative: 102
Best Local Similarity: 22.80% Mismatches: 162
Query Match: 14.42% Indels: 305
DB: 24 Gaps: 12

US-10-051-909-32 (1-800) x ABL41880 (1-1518)

QY 55 LysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGly 74
Db 19 AAAAAGCAGTCAATATATAGCTTTATTTTCGAGACTCTTGAGGCGGCTTATATGCG 78
QY 75 TrpAspAsnAlaThrIleAlaAlaValLeuYrIleLysLySGluPheGlnLeuGln 94


```
Db 79 TATGATACCGAGTGAATTCCTGGAGCTATTTTATTATGAAAAAGAGTTAGCCTTA--- 135
QY 95 AsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 136 ---AACGCGTTTACAGAAAGTCTGTGTGTGACCTCTGCTGTGGCGCATATTTGGGC 192
QY 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSer 134
Db 193 TCAGAGCGCGCCGCAAGCTGACTGACCGCTTTCGGAAGAAAAAAGCAATTTAGCAGACC 252
QY 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
Db 253 GCGCTGCTGTTTGTATAGCGCGTCTGTGTGGCACTGGCCCAATACAGAGTCATG 312
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 313 GTGCTGTTTCGATCATTTTGGGACTTGACATCGAACAATCGTACCCCTT 372
QY 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 373 TATTATCTGAACCTGGCGCCAAACATAAACCGCGCGCTGTCACTCACTGAATCAGCTG 432
QY 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 433 ATGATCACGTCGCGCATCTTCTTTCTTAC-----ATTGTCAATTACATATTTGCCGAT 486
QY 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db 487 GCCGAAGCGTGGCGCTGATGCTTGATTTGGCTGTGCTGCTCATTTGCTCTTCTT--- 543
QY 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMetAla 254
Db 544 ATTGGCATTTTGTATTATGCCGGAAGACCCGCGCTGCTGTTCACGAATGGCGAAGAAAGC 603
QY 255 GluAlaIleValSerValLeuGlnIleValArgGlyIleAspAspValSerGlyLeuLeuSer 274
Db 604 AAAGCGAAGAAATTTCTGAAAAAATGCGTGGCACAAGATATT----- 648
QY 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluIleTyrIleIleGly 294
Db 648 ----- 648
QY 295 ProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIleGluIleThrLeu 314
Db 648 ----- 648
QY 315 TyrGlyProGluGluGlnSerTrpIleAlaArgProSerIleGlyProIleMetLeu 334
Db 648 ----- 648
QY 335 GlySerValLeuSerIleValSerArgHisGlySerMetValAsnGlnSerValProLeu 354
Db 648 ----- 648
QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnIleGlyGly 374
Db 648 ----- 648
QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
Db 648 ----- 648
QY 395 AlaIleAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluIleTyrAlaSer 414
Db 649 -----GATCAGGAA----- 657
QY 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db 657 ----- 657
QY 435 ThrGlyAlaGluGlyIleAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
Db 658 -----ATACAT----- 663
```

```
QY 455 ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
Db 664 -----GATATA----- 669
QY 475 GlyTrpGlnLeuAlaTrpIleTrpSerGluIleGlyGluGlyIleAsnGlyArgIleGlyGly 494
Db 670 -----AAAGAACCGGAAAGCAGATGAAGCGGCTGTGAAGAG--- 708
QY 495 GlyPheIleArgValTyrLeuHisGlnGluGlyValProGlySerArgArgIleSerIle 514
Db 708 ----- 708
QY 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534
Db 708 ----- 708
QY 535 LeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAspAla 554
Db 708 ----- 708
QY 555 AlaMetValHisProSerGluValAlaAlaIleGlySerArgTrpIleAspLeuPheGlu 574
Db 709 -----CTGTTTCGAT 717
QY 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594
Db 718 CCATGGGTGCGCCAGCGCTTATTGACAGTTTGGGACTCGCTTTTGTGCAATTTATC 777
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
Db 778 GGAACGAATACGATCATCTACTATGCGCCAAAGACCTTACAACGTCGGA----- 828
QY 615 IleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeu-----Ile 630
Db 829 -----TTCCGA-----AACTCCGCTTCGATTTTAGGCACCGTCGGAATC 867
QY 631 SerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAsp 650
Db 868 GGCACAGTCAATGTTCTCATGACATTA-----GTAGCGATTAAATCATCGAC 915
QY 651 LeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670
Db 916 AAGATTGGAAGAAAGCGCTTACTGCTATTCGGAATCGCGCATGCTGATCAGCTTGATC 975
QY 671 IleLeuValIleValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
Db 976 GTTCTCGCTTAGTAATCTCTTTTCAATAACACTCCGCGTCCCTCATGACGACCGTC 1035
QY 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
Db 1036 ATTTGTTTAGCGCGTGTATTATCGTGTCTTGGCGGTCAAGTGGGACCGGTTGTGGGTG 1095
QY 711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730
Db 1096 ATGCTTCCTGAATTGTTCCCGCTTCACGTCAGAGGAATCGGACCGGTGTTTCGACCTTA 1155
QY 731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
Db 1156 ATGTTGACAGTTGGGACACTGATTTTCATTAACCTATCCAAATATTAAAGAGCGATC 1215
QY 751 GlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheVal 770
Db 1216 GGAATCAGTATTATTATCTGATTATGCGCGCATCGGTATCATGCGCTTCTTATTGTC 1275
QY 771 PheLeuIleValProGluThrIleGlyMetProLeuGluValIleThrGlu 787
Db 1276 CGATTTAAGTGACAGACAGACAAGGAGAGAGCCTTGAAGAATTGAGCAG 1326
RESULT 14
ABK51964
ID ABK51964 standard; cDNA; 870 BP.
XX
```

AC ABK51964;
XX 27-AUG-2002 (first entry)
XX Rice cDNA clone rl56.pk0003.d5 encoding sugar transport protein.
DE
XX
XX Rice; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant; clone rl56.pk0003.d5; gene; ss.
XX
OS Oryza sativa.
XX
XX
FH Key Location/Qualifiers
FT CDS 8..403
FT /*tag= a
FT /partial
FT /product= "Portion of a rice Arabidopsis
FT thaliana-like sugar transport protein"
FT /note= "This sequence lacks a start codon"

XX US6383776-B1
XX 07-MAY-2002.
XX
XX 14-APR-1999; 99US-0291922.
XX
XX 24-APR-1998; 98US-083044P.
XX
XX (DUPO) DU FONT DE NEMOURS & CO E I.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-451386/48.
DR P-PSDB; AAU97203.

XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
PT
XX
PS Claim 3; Column 33-36; 54pp; English.

XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence encodes
CC a portion of a rice Arabidopsis thaliana-like sugar transport protein.
XX

XX Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;

Alignment Scores:
Pred. No.: 1.03e-39 Length: 870
Score: 562.00 Matches: 103
Percent Similarity: 93.75% Conservative: 17
Best Local Similarity: 80.47% Mismatches: 8
Query Match: 13.79% Indels: 0
DB: 24 Gaps: 0

US-10-051-909-32 (1-800) x ABK51964 (1-870)

QY 669 leuValileuValSerAnleuIleAspLeuGlyThrleuAlaHisAlaLeu 688
Db 8 GTTCTAACCTTGATTCGTCAATATCTCGATGTGGGACCATGTTCATGCCCTCACTG 67
QY 689 SerThrValSerValileValIleValTyrPheCysSpheValMetGlyPheGlyProIlePro 708
Db 68 TCCACAGTCAGTGTCACTACTCTCTGCTTCTTGTCAATGGGGTTCGGGCTTATTC 127
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728

Db 128 AACATTTCTGTGACGATTTTCCCGACCACCGTTCGTGCATCTGCATGCCATCTGT 187
QY 729 AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748
Db 188 GCCCTAACATTTCTGATCGGTGATATCATGTGCATACACCCCTCCCGTATGCTCAAC 247
QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal 768
Db 248 GCCATTGACTCGCTGAGTGTTCGAATCTACGAGTGTCTGCATACTGCTTCTCTG 307
QY 769 pheValPheLeuValProGluThrIleGlyMetProLeuGluValIleThrGluPhe 788
Db 308 TTTGTCTTCATGAAGTCCCGAGACAAAGGCATGCCCTTGAAGTCATCACCGAGTTC 367
QY 789 pheAlaValGlyAlaIleGlnAla 796
Db 368 TTCTCTGTGAGCAAGCAGGCC 391

RESULT 15
ABX93200
ID ABX93200 standard; cDNA; 870 BP.
XX
XX ABX93200;
AC
XX
DT 29-MAY-2003 (first entry)
XX
DE cDNA encoding rice sugar transport protein #2.

XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Oryza sativa.

XX US2002178468-A1.
XX
XX 28-NOV-2002.

XX 17-JAN-2002; 2002US-0051902.
XX
XX 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.

XX (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.

XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;

XX WPI; 2003-340957/32.
DR P-PSDB; ABU08328.

XX Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
PT
XX
PS Claim 2; Page 18; 56pp; English.

XX The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX

Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;

Alignment Scores:

Pred. No.:	1.03e-39	Length:	870
Score:	562.00	Matches:	103
Percent Similarity:	93.75%	Conservative:	17
Best Local Similarity:	80.47%	Mismatches:	8
Query Match:	13.79%	Indels:	0
DB:	25	Gaps:	0

US-10-051-909-32 (1-800) x ABX93200 (1-870)

QY	669	LeuValIleLeuValIleSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeu	688
Db	8	GTCTTAACCTTGATTCGTGCAATATCTGTGATGTGGGACCATGTCATGCCCTCACTG	67
QY	689	SerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro	708
Db	68	TCCACAGTCAGTGTCACTACTCTTCTGCTTCTTGTCAATGGGGTTCCGGCCTATTCCA	127
QY	709	AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys	728
Db	128	AACATTCTCTGTGACAGATTTCGCCGACCACCGTTCGTGCAATCGCATAGCCATCTGT	187
QY	729	AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn	748
Db	188	GCCCTAACATTCGTGATCGGTGATATCATTTGACATACACCCCTCCCGTGTGCTCAAC	247
QY	749	AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal	768
Db	248	GCCATTGCACTCGCTGAGTGTTCGAATCTACGCAAGTGTCTGCATACTGGCTTTCCTG	307
QY	769	PheValPheLeuValProGluThrLysGlyMetProLeuGluValIleThrGluPhe	788
Db	308	TTGTCTTCATGAAGGTGCCGAGACAAAGGCGATGCTCTTGAAGTCATCACCAGTTTC	367
QY	789	PheAlaValGlyAlaLysGlnAla	796
Db	368	TTCTCTGTGAGCAAGCAGGCC	391

Search completed: January 5, 2004, 19:18:26
Job time : 617 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 18:59:34 ; Search time 7686 Seconds
(without alignments)
4258.093 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPLDRR.....PLEVITRFFAVGAKQAAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10051909/runat_05012004_094741_22306/app_query.fasta_1.967
-DB=GenEmbl -QFMT=fastap -SUFFIX=xge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1_1_4944 @runat_05012004_094741_22306 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3517	86.3	2824	6 AR208565	AR208565 Sequence
2	3155	77.4	2665	8 AY165599	AY165599 Saccharum
C	3052.5	74.9	137140	2 AP005756	AP005756 Oryza sat
	2808	68.9	2614	8 HVU534445	AU534445 Hordeum v
	2674	65.6	2601	6 AR208568	AR208568 Sequence
	2657	65.2	2570	8 AY094465	AY094465 Arabidops
6	2657	63.4	2190	6 AX506620	AX506620 Sequence
7	2584	63.4	2190	8 ATH532570	AX532570 Arabidops
8	2584	63.4	2190	8 AC073166	AC073166 Oryza sat
9	2514	61.7	300957	8 AE017116	AE017116 Oryza sat
10	2514	61.7	300957	8 ATE23E12	AL022604 Arabidops
11	2502	61.4	86710	8 ATCHRIV83	AL161587 Arabidops
C	12	2502	58.1	8 AC136843	AC136843 Oryza sat
	13	2368	58.1	8 AC136843	AC136843 Oryza sat
	14	2275	55.8	6 AX412656	AX412656 Sequence
	15	2275	55.8	6 AX507559	AX507559 Sequence
16	2245	55.1	2190	8 ATH532571	AJ532571 Arabidops
17	2217	54.4	2426	8 ATSGUTRPR	Z50752 A.thaliana
18	2191.5	53.8	2516	8 HVU534446	AJ534446 Hordeum v
19	2144	52.6	103192	8 AC007369	AC007369 Arabidops
20	2144	52.6	143879	8 AC069251	AC069251 Genomic s
21	2128	52.2	114918	8 AP004945	AP004945 Lotus jap
22	2117	52.0	94349	8 ATF26013	AL133452 Arabidops
C	23	2093.5	51.4	2 AC121239	AC121239 Medicago
	24	1968	48.3	2 AP004082	AP004082 Oryza sat
	25	1959	48.1	2 AC126786	AC126786 Medicago
	26	1927.5	47.3	2 AC144482	AC144482 Medicago
27	1670.5	41.0	1692	6 AR208569	AR208569 Sequence
28	1557.5	38.2	87592	8 AC099739	AC099739 Oryza sat
29	1557.5	38.2	127098	8 AC144426	AC144426 Oryza sat
30	1557.5	38.2	154128	8 AP000615	AP000615 Oryza sat
31	1424	34.9	1487	6 AR208571	AR208571 Sequence
C	32	1159.5	28.5	8 AF128457	AF128457 Oryza sat
	33	1140	28.0	8 AF119222	AF119222 Oryza sat
	34	1140	28.0	8 AF161269	AF161269 Oryza sat
	35	1055.5	25.9	2 AC127019	AC127019 Medicago
C	973	23.9	106246	8 AF528565	AF528565 Zea mays
C	973	23.9	346296	8 AF090447	AF090447 Zea mays
C	971	23.8	108553	8 AF527809	AF527809 Sorghum b
C	969.5	23.8	131457	2 AC144718	AC144718 Zea mays
40	943	23.1	1009	6 AR208572	AR208572 Sequence
41	587.5	14.4	1518	6 AX353444	AX353444 Sequence
42	587.5	14.4	8189	1 BSZ92954	Z92954 B.subtilis
C	587.5	14.4	212610	1 BSUB0019	Z99122 Bacillus su
44	562	13.8	870	6 AR208567	AR208567 Sequence
45	539	13.2	1395	6 AX432492	AX432492 Sequence

RESULT 1

ALIGNMENTS

AR208565 AR208565 2824 bp DNA linear PAT 20-JUN-2002
LOCUS AR208565
DEFINITION Sequence 1 from patent US 6383776.
ACCESSION AR208565
VERSION AR208565.1 GI:21509752
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2824)
AUTHORS Allen, S.M., Hitz, W.D., Kinney, A.J. and Tingey, S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 1 07-MAY-2002;
FEATURES location/Qualifiers
source 1. 2824
/organism="unknown"
BASE COUNT 644 a 649 c 745 g 775 t 11 others
ORIGIN

Appl

Alignment Scores:
Pred. No.: 2.13e-225 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: Gaps: 2
US-10-051-909-32 (1-800) x AR208565 (1-2824)
QY 56 MeuSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db 238 ATGGGGGGCGCGGTGATGTCGCCATCGCGGCTCTATCGGCACTTGCTGCAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleIleGlyGluPheGlnLeuGlnAsn 95
Db 298 GACAAATGCCAAATGCTGAGCGCTCTGTACATAAAGAGGAATTCAACCTGCAGAC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTCTGATCGAAGGCTCATGCTCGCATATGTTCTCATTTGGGCAACAGTCATCACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCCGGGGCCAGGCTGATGCTGTTGTAAGGCCCATGCTGTCGCTCGCT 477
QY 136 IleuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 155
Db 478 GTCTCTACTTCTGTCAGTGGGCTGATGCTTG3GCGCAATGTGTACATCTGTCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 538 CTGGCAAGGCTCATGTATGGGTTGGGTTGCGGTCACTTGTCTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAACCTGCACCGCACAGANATTCTTGGGCTGNTNGAACACAGTTGCCGAGTTC 657
QY 195 SerGly--SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATGGGGTCAAGGAGGATGTTCTCTCTACTGATGTTGGATGTTGGATGTTCCCTCATG 717
QY 214 ProSerProAspTyrPargIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 718 CCCAACCCTGATGGAGGCTCATGTTGAGTTCTGTGATCCCGTCACTTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrPheValSerIleGlyArgMet 253
Db 778 GGACTGACTGTCTTCTACTTGTCTGATACCAAGTGGCTTGTNAGCAAGGAAGATG 837
QY 254 AlaGluAlaIleValLeuGlnIleValLeuArgGlyIleAspAspValSerGlyGluLeu 273
Db 838 GCGGAGGCGAAGAGAGTGTGCAAGGCTGCGGGAAGAGAGATGTCTCANGGAGANG 897

QY 274 SerLeuLeuGlnGlyLeuGluValGlyGlyAspThrSerIleGlnGlyTyrIleIle 293
Db 898 GCTCTTCTAGTGAAGGTTTGGGGTCCGTAAGATACACGATATTNAGAGTACATCAT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIleGlnIleThr 313
Db 958 GGAACCTGCCACCGAGGACCGCATGATCTTGTACTGACCGGTGATTAAGACAATCACA 1017
QY 314 LeuTyrGlyProGlnGluGlyGlnSerTyrIleAlaArgProSerIleGlyProIleMet 333
Db 1018 CTTTATGGGCTGAAGAGGCCAGTCATGATGCTCGACCTTCTAAGGACCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 1078 CTTGAAGTGTGCTTCTCTGTCATCTCGTCAATGGAGCATGGTGAACCAAGTGTACCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGATCCGATGTGACACTTTTGGTAGTCCATGAGAATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGTATGAGGACATTTGTTCCAACTTTGGAAGTATGTTCAAGTGTACAGATCAG 1257
QY 394 HisAlaIleAsnGlnGlnTyrPaspGluGlnLeuHisArgAspAspGluTyrAla 413
Db 1258 CATGCCAAATGACAGTGGATGAAGAGAACTTCTATAGGATGACGAGAGTACGCA 1317
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGTGACAGAGGTGACTATGAGACAAATCTCCATAGCCCATGCTGTCCAGGACG 1377
QY 434 AlaThrGlyAlaGlnGlyIleValHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGCGAAGGAGAACATTTGTCACATGTCACCCGTGAAGTCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAAGCAAGGCTCTTAGGGAGGAGTGAAGATGTTGAGCAGCATGATACGGT 1497
QY 474 GlyGlyTyrGlnLeuAlaTyrPlySerTrpSerGluGlyGluAsnGlyArgGlyGlu 493
Db 1498 GGGGATGGCAGCTTCTTGAATGTGTGACAGAAAGGTGAGATGTGAGAAAGAA 1557
QY 494 GlyGlyPheIleArgValTyrLeuHisGlnGlnGlyValProGlySerArgGlySer 513
Db 1558 GGTGTTCAAAAGAGTCTACTTGCACCAAGGAGTCTCGGCTCAAGAGGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGlyGlySerGluPheValHisAlaAla 533
Db 1618 ATTGTTCACTTCCCGGTGGTGGAGTGTGAGGTAAGTGTGATGATGCTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTAGTAGTCAAGTCAAGCACTTTCTCAAAAGGCTTGTGAACCAAGCATGTCAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaIleGlySerArgTyrPlyAspLeuPhe 573
Db 1738 GCTGCATGTTCACCCATCTGAGGTAGCTGCCAAAGTTCAAGTGAAGATTTGTTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePhe 593
Db 1798 GAACCTGAGTGAAGGCTGCCCTGTAGTGGGTGTGAATTTCAGATCTTCAACAGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
Db 1858 GCTGAATAAAGGTGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGTGTGCA 1917
QY 614 ValIleLeuSerIlePheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATTTCTTCCAAATTTGCTCAGCTCGGCATCAGCATCCATCTGATCAGTTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653

Db 1978 ACTACCTTAATGCTTCCTTGCACTTGCTTGCCATGCTTATGATGATCTTCCGA 2037

QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673

Db 2038 AGAAGTTTGTGCTGACGACATCCATCTGATGACATCTCTAGTATCTCGTT 2097

QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693

Db 2098 GTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157

QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713

Db 2158 ATCGTCTACTTCTGCTGCTTCGTATGGGATTGGTCCCATCCCAACATTTTATGCA 2217

QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733

Db 2218 GAGATCTTCCCAACCAAGGTTGCGCTCTGTATGTCATTTGGCTTACATCTGG 2277

QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753

Db 2278 ATCGGAGATATCATCGTCACTACAGCCTTCTGTGATGCTGAATGCTATTGGACTGGC 2337

QY 754 GlyValPheSerIleTyrAlaValAlCysLeuIleSerPheValPheValPheLeuLys 773

Db 2338 GGTGTTTTCAGCATATATGCACTGCTGATTTCTTTGTGTTGTTGTTGTTGTTAAG 2397

QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793

Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGTTATTACGAATCTTTGCAGTGTGCG 2457

QY 794 LysGlnAlaAlaAlaLysAla 800

Db 2458 AAGCAAGCGGCTGCMAAGCC 2478

RESULT 2

LOCUS AY165599 2665 bp mRNA linear PLN 07-JUN-2003

DEFINITION Saccharum hybrid cultivar putative sugar transporter type 2a mRNA, complete cds.

ACCESSION AY165599

VERSION AY165599.1 GI:31505503

KEYWORDS

SOURCE Saccharum hybrid cultivar (sugarcane)

ORGANISM Saccharum hybrid cultivar

REFERENCE 1 (bases 1 to 2665)

AUTHORS Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.

TITLE Identification of a novel sugar transporter homologue strongly expressed in maturing stem vascular tissues of sugarcane by expressed sequence tag and microarray analysis

JOURNAL Plant Mol. Biol. 52, 371-386 (2003)

2 (bases 1 to 2665)

CASU,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.

TITLE Direct Submission

JOURNAL Submitted (17-OCT-2002) Division of Plant Industry, Commonwealth Scientific and Industrial Organisation, 120 Meiers Rd, Indooroopilly, QLD 4068, Australia

FEATURES

source

1..2665

/organism="Saccharum hybrid cultivar"

/mol_type="mRNA"

/db_xref="taxon:128810"

97..2334

/codon_start=1

/product="putative sugar transporter type 2a"

/protein_id="AA037640.1"

/db_xref="GI:31505504"

/translation="MSGALVAIVASIGNLLQGMNATISALLYIKKEFKLESEPTV

EGLIVAMSLIGATITTFSGPVSQWIGRLPMLILSSVLYPVSSLIMLSPNVYLLA

RLINGFGVLAIVLPLVLISETAPEIRGLINTLPQFSGSGGMLSYCMVFGMSILPS

PDWRIMGLVALPSLFFGLTIFYLPESPRVLVSKGRMAEKVYLQRLRGKDDVTGEM

ALVEGLEVGDDTFIEEYIIGPAIEPADDHVADDERITILFGEPEGQSWIARPSKGS

SMLSVLSIGSRHSGSMVNQNLPLTDPVTLFGSVHENISLAGSMRRIVPSKFDMSIS

ITDQHPKTEQWDEENLHRDNEEYATDGAASDYENNVSPLSRQNTSABGKIAHGH

RGSALGLRRSLSDGEAEATSTGICGWQLAWKWSEREDEGCKESFKRIYIHOEG

VAGSRGVSLSLPGGDASEGCKFIHAALVSQALYPRDITEQMAGPATMHPSEAA

AKVPSMRDLPFEGVRALLVGIQILOQFAGINGVLVYTPQIMEQAGAVALLSNLGL

SSASASILISSVTALMLPSIGLAMELMDVSGRRFLLSITPVLIASLIVLVSNVIE

LGTVHAVLSTISVITYLCCFKMGFPIPNILCAEFPTVRVIGICAIICALIFWVGI

IVTVSLPWLNAIGLEGVFGIYAVACAIAFVYVLKVPETKGMPLVITEFFAVGAKQ

AVAKA"

misc_feature 97..156

/note="putative; transmembrane-region site"

misc_feature 115..774

/note="pfam00083; Region: sugar transporter"

misc_feature 238..297

/note="putative; transmembrane-region site"

misc_feature 313..372

/note="putative; transmembrane-region site"

misc_feature 406..462

/note="putative; transmembrane-region site"

misc_feature 511..570

/note="putative; transmembrane-region site"

misc_feature 589..654

/note="putative; transmembrane-region site"

misc_feature 658..1794

/note="Region: large cytoplasmic loop"

misc_feature 1624..2298

/note="pfam00083; Region: sugar transporter"

misc_feature 1672..1731

/note="putative; transmembrane-region site"

misc_feature 1813..1869

/note="putative; transmembrane-region site"

misc_feature 1897..1953

/note="putative; transmembrane-region site"

misc_feature 1969..2022

/note="putative; transmembrane-region site"

misc_feature 2089..2151

/note="putative; transmembrane-region site"

misc_feature 2191..2247

/note="putative; transmembrane-region site"

BASE COUNT 650 a 561 c 691 g 763 t

ORIGIN

Alignment Scores:

Pred. No.: 2.93e-201 Length: 2665

Score: 3155.00 Conserves: 601

Percent Similarity: 89.56% Mismatches: 68

Best Local Similarity: 80.46% Indels: 0

Query Match: 77.42% Gaps: 0

DB: 8

US-10-051-909-32 (1-800) x AY165599 (1-2665)

QY 54 AspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGln 73

Db 91 GAAACATGTGCGGAGCTGCTTGTGTCATATGATGCTTGCATCGCAATCTGCTCAG 150

QY 74 GlyTyrPaspAsnAlaThrIleAlaAlaValLeuTyrIleLysGlnPheGlnLeu 93

Db 151 GGGTGGACATGCCCATCTCAGCTGCTCTATTATATATAAGAGCAATTCAAACCTA 210

QY 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113

Db 211 GAAAGTAGCCTACTGTGAGGGGCTGATTTGGCCATGTCACTAATTGGAGCAACTATC 270

QY 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133

Db 271 ATTACACATTTCTTGACACAGTATCAGACTGGAATGGCCGCTCCCTATGCTCATCTC 330

QY 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153

Db	331	TCCTCAGTTCTGTACTTTGTGTAGACACCTCATTTATGCTATGTTCCCGAATGTCATATGTA	390
QY	154	LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro	173
Db	391	CTACTGCTGGCACGCTTTATTAATGCAATTCGGTGTCCGCTTGCTGCTCACTCTGTTCCT	450
QY	174	LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln	193
Db	451	CTGTACATTTTCAGAAACAGCTCTCCACAGATAAGAGTTTGTGTAATACACTGCCACAG	510
QY	194	PheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer	213
Db	511	TTTCAGTGGGTCAGGGGGATGTTTTTGTCTGCTACTGCATGTTGTTCCGGATGTCACTGCTG	570
QY	214	ProSerProAspTyrPargIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe	233
Db	571	CCATTCACCTGATTTGAGGATTAATGCTTGGGGTTCTCCGGTTACTTCATTGTTTCTTT	630
QY	234	GlyLeuThrIlePheTyrLeuProGlySerProArgTyrPheValSerIlyArgMet	253
Db	631	GGTTTGACAATATTTTATCTGCTCAATCACCAAGATGGCTGTTCAGCAAGGGCGAATG	690
QY	254	AlaGluAlaIlyIysValLeuGlnIlyLeuArgGlyIlyAspAspValSerGlyGluLeu	273
Db	691	GCACAGAGCAAGAAGAGTGTTGCACAAGATTACGTGGCAAGAGTACGTCACAGCGCAATG	750
QY	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle	293
Db	751	GCCTCTCTTGTGGAAGGGTTGAGGTGCGAGGGGATACCTTCATTGAAGATACATCATT	810
QY	294	GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIlyGluGlnIleThr	313
Db	811	GGTCCTGCTATCGAACCAAGCTGATGACCAAGTCGCTGATGGTATAGAGAAGCAATAATA	870
QY	314	LeuTyrGlyProGluGluGlnSerTyrIleAlaArgProSerIlyGlyProIleMet	333
Db	871	CTATTGGGGCCTGAAGAGGGCCATCATGATAGCTCGGGCTTCCAAGGGTCCAGCATG	930
QY	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	931	CTTGGAGTGTCTTCTTATCGGATCTCGTCATGGCAGCATGTGAAACCAAAATTACCC	990
QY	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	991	CTTACGGATCCTATTGTACAGCTTTTGGAGTGTCCATGAGAACATCTCTCGCTGGA	1050
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1051	GGAGTATGAGGAGATAAGTGCCTTAATTTGACAGTATGATCAGTATCACAGATCAG	1110
QY	394	HisAlaIlyAsnGluGlnTyrAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1111	CAGCCCAAGACCGAGCAGATGGGATGAGAGAACTCCATAGGGATAATGAGAGTATGCA	1170
QY	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1171	ACTGATGTGCTGCAAGTGACTATGAACAAATGTACACAGCCCTGCTGTCCCGACAG	1230
QY	434	AlaThrGlyAlaGluGlyIlyAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453
Db	1231	AATACAAGTCCGAGGAGGAGGCAATTGCACATCACGGGCACTGGAAGTCTTTGGGT	1290
QY	454	MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1291	TTGAGAGAAGAAAGACCTGTCCGATGAGGGGTGTGAAGCAGCAACCACTGGCATTTGGT	1350
QY	474	GlyGlyTyrPglLeuAlaTyrIlySerTyrSerGluIlyGlyGluGluAsnGlyArgIlyGlu	493
Db	1351	GGCGGATGGCAGCTCGCCTGGAATGTGTACAGCGAAGAGGTGAGACCGTAAGAAGGAA	1410
QY	494	GlyGlyPheIlySerGluValTyrIleHisGlnGluGlyValProGlySerArgArgGlySer	513

Db	1411	GGGAGTTTCAAAAGAAATCTACTTACACCAAGAGGAGTGTGCTGCTCAAGAGGGGATCT	1470
QY	514	ILEValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1471	GTTGCTCACTTCTCGTGGAGGTGATGCGCTGAAAGGTGGTAAGTTCATACATGCTGCT	1530
QY	534	AlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAsp	553
Db	1531	GCTCTGGTCAGCCAGTCAGCTCTTTACCCGAGGATATTACAGAACAGCGCATGGCCGGT	1590
QY	554	AlaAlaMetValHisProSerGluValAlaAlaIleGlySerArgTyrIleAspLeuPhe	573
Db	1591	CCAGCTACAATGACCCATCAGAGGAGCAGCTGCCAAAGTTCCAAGCTGGAGAGACCTGTTT	1650
QY	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePhe	593
Db	1651	GAACTGTGTGAGAGCTGCCCTGTGGTGGTATGGAATTGAGATTCATTCACACAGTTT	1710
QY	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1711	GCGGGAATAAATGGGGTCTCTACTACACTCCGCAAAATTATGAGAGCAAGCTGGGTGCA	1770
QY	614	ValIleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633
Db	1771	GTTCTTATTCCAACTCTTGCTCTCAGTTCGGCATCAGCATCCATCCTAATTAGTTCGTCTC	1830
QY	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1831	ACTGCCCTACTTATGCTCCCTAGCATTGGTTAGCCATGAGACTTATGAGCGTATCTGGA	1890
QY	654	ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	1891	AGAAGTTTCTGCTGCTAAGCACAAATTCAGTCTGATAGCATCTTAAATGTCCTGTTT	1950
QY	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	1951	GTTGCTAATGTTATGCAAGTTGGGTACAGTGGTCCATGCTGCTCTCCACAAATTAGTGTC	2010
QY	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2011	ATCACCTACTTATGCTGCTTCAAATATGGGCTTGGCCCCCATTTCCCAACATTTATGTGCA	2070
QY	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr	733
Db	2071	GAGTTTTCACACAGGGTTCGTGGCATTCGCAATTCGCAATCTGTGCTTGAATTTTGG	2130
QY	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2131	GTTGAGACATTATTCACATACATACAGTCTTCCGTGATGCTGAATGCTATTGGTCTTGA	2190
QY	754	GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLys	773
Db	2191	GGCGTGTGTCATATACGCAAGTTCCTGTGCAATTCCTTTGTGTTGTCTAATCTTAAG	2250
QY	774	ValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793
Db	2251	GTTCCGGAGACAAAGGCGCATGCCCCCTTGAGGTATCATCTGAGTTCCTTGGCAGTTGGTGCA	2310
QY	794	LysGlnAlaAlaAlaLysAla 800	
Db	2311	AAGCAAGCGGTGCAAGGCC 2331	

RESULT		3			
AP005756/c					
LOCUS					
DEFINITION	AP005756	137140 bp	DNA	linear	HTG 19-SEP-2002
	Oryza sativa	(japonica cultivar-group)		chromosome 2	clone
	OSJNB0035N08,	*** SEQUENCING IN PROGRESS ***.			.
ACCESSION	AP005756				
VERSION	AP005756.1	GI:23200615			
KEYWORDS	HTG; HTGS PHASE2.				
SOURCE	Oryza sativa	(japonica cultivar-group)			
ORGANISM	Oryza sativa	(japonica cultivar-group)			
	Eukaryota; Viridiplantae;	Streptophyta; Embryophyta; Tracheophyta;			

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

1
REFERENCE
AUTHORS
TITLE
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC clone:OSJNB0035N08
JOURNAL
REFERENCE
TITLE
Published Only in Database (2002)
2 (bases 1 to 137140)
Sasaki,T., Matsumoto,T. and Katayose,Y.
Direct Submission
JOURNAL
Submitted (18-SEP-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. 137140
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSJNB0035N08"
BASE COUNT 40995 a 27904 c 27671 g 40414 t 156 others
ORIGIN
Alignment Scores:
Pred. No.: 6.23e-192 Length: 137140
Score: 3052.50 Matches: 672
Percent Similarity: 50.81% Conservative: 46
Best Local Similarity: 47.56% Mismatches: 70
Query Match: 74.91% Indels: 625
DB: 2 Gaps: 7
US-10-051-909-32 (1-800) x AP005756 (1-137140)
QY 3 SerGlySerTrpLeuAlaValGlnThrProAspLeuAspArgGluArg 22
Db 88617 GCTGCCTCCTGGCTTCCATTGAGCCGCTTCGATTGATCTGGAGCCGAGACACC 88558
QY 23 LeuLeuProSer----- 26
Db 88557 CCGCTTCCTCCTCCCTCGGTTCTTGCGCTTCTTTGGAGAGTCCGCTG 88498
QY 27 -----ValValLeuAlaLeuProGlyPro 34
Db 88497 ATGAGTTCTGCAAGAATTTGGGAGTCTGTGTTGTTGTTATGTGATGCTGCTCG 88438
QY 35 LeuProProAlaSer-----CysSerSerGlnGluProValThrSer 48
Db 88437 GGTGTGTTCAATTTCTGCTGCTGATTCTTGATGACAGGAGTCGAGAGGGGTG----- 88384
QY 49 AspAspIleLeuGluAspIysMetSerGlyAlaValLeuValAlaIleValAlaSerIle 68
Db 88383 -----AAGATGTCGGGTGCTGCTCTGTCGCAATCGCGGCTTCATC 88342
QY 69 GlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAla----- 81
Db 88341 GGCACCTGCTGCAGGGATGGACAATGCCACCATTTGAGGTAAAGGTTTACTTCCC 88282
QY 81 ----- 81
Db 88281 GTGTGCTAGATTGTTATTTGTTGAGAGTTTGTGACTAGAAATATGAAACAGC 88222

QY 81 ----- 81
Db 88221 TGTCAATTTTCTTAATTATCAGTGTGGTACCTTCAATTTTCAGATTATAGA 88162
QY 81 ----- 81
Db 88161 CTGCTTAGATCAGATGTTTGTGAGTGTAAATTATGCTCTTCTAGTAACATGCTGTG 88102
QY 81 ----- 81
Db 88101 TTGTTGTTAGTCTTAAGACCAATTGGCAAAATTGTTATCAGAAACTGGATCGCATG 88042
QY 81 ----- 81
Db 88041 ATATTAATAGTCTTAATTTTCTAGATCTGTAATACTTATTTGAAGAAAGATGTT 87982
QY 81 ----- 81
Db 87981 CTAATCATATGTTAGCTATTGTGTGTCTGTGAATCATTACTGTTGCCATTCTT 87922
QY 81 ----- 81
Db 87921 TTCTGAACACCTGTAAGATGTTCTATCCACAAGAAAGTGAATGTTAACAGGCAC 87862
QY 82 -----AlaAlaValLeuTyrIleLysLysGluPheGlnLeuAsn 95
Db 87861 AATTATTAACCTTCTCAGGTGCTGTACTATACATAAAGAAAGAACTAGAAAGT 87802
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 87801 GAGCCCACTGTGAGGGGCTAATCGTGCCATGTCATGTTGTTGCAACCATCATCACA 87742
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSerSer 135
Db 87741 ACATTCAGAGGCGCGTATCAGACTGATCGCGCGCGCTATGCTCATCTCTTCA 87682
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
Db 87681 ATTCTTACTTCTTCAGACAGCCTCATCATGTTATGTTCCCGAATGTTATGTTACTG 87622
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 87621 CTCGACGCTTATAGATGATTCGCGATCGGCTGCTGCACACTGTATCCTTTGTAC 87562
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 87561 ATCTCAGACAGACGCTCCTCAGAGATCAGGGGTTGCTGAATACACTGCCACAGTTCA 87502
QY 196 GlySerGlyIleMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 87501 GGGTCAGAGGAGATGTTCTTGTCTCCTACTGATGTTGGATGTCATGTTGCCGTCA 87442
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 87441 CCTGACTGAGAAATTATGCTTGTGTTCTCGCAATCCCTCGTTGTTTCTTCGGAATG 87382
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 87381 ACAATATTCTATCTGCCAAGATCACCAGATGCTTGTACGAAAGGCGGATGGCTGAG 87322
QY 256 AlaLysValLeuGlnLysLeuArgGlyLysAspValSer----- 270
Db 87321 GCAAGAAGGATATTACAAAAATTACGTGGAGAGAGATGTCTCAGGTATGTTTGTCTGA 87262
QY 270 ----- 270
Db 87261 AAATTTGATGTTCAATTCACCATATTTCCAGCTGACCAAGTTAAACAATTCTCCC 87202
QY 270 ----- 270
Db 87201 AGATTCTAAACTCATGCTTAATACTAGAAACATGTGTGGCATTTGTAAGTTTTC 87142

QY 270 ----- 270
Db 87141 TTGACAGCACTTAACATATGCAAAATAGTATTTTGAAAAAAGGTAAGAAA 87082
QY 270 ----- 270
Db 87081 TCAAAATCTATTAGCCCTTCTATCGGTGCTTTTGA AAAAGCTTATGAATCTTCAAG 87022
QY 270 ----- 270
Db 87021 TAAAAATTACTAAGTACACATACTACACTAATTCATTTCCATAGAGGCAGAAACATGG 86962
QY 270 ----- 270
Db 86961 AATGTACAATCGTTATCTCAACAGGTAGAGCCCAAATGAGGTGTCCCAACTTTTGTCTGA 86902
QY 270 ----- 270
Db 86901 GTCCTAATCTGAGAAAAATAATCAGGTGATCTAGATAGACATCGCCCTCTTTT TTTT 86842
QY 270 ----- 270
Db 86841 TTTTGTAAAGGTAAGTCTAGAAATTTGGGGGGTGGACCTGCTGTATATGTGATGTG 86782
QY 270 ----- 270
Db 86781 GGCCGACCCACGGGCTCTTATCTACTAGTTAACTAAATGAACATATATTTGCTTTTG 86722
QY 270 ----- 270
Db 86721 AGATTTTAAAAAGTAAATTAAGTCCATCTAGTTGACCTAGAGTGGGTCAAGGAGCCCAAC 86662
QY 270 ----- 270
Db 86661 AGCGCCCATGCTTAACAACAGTATGTTCAAGAAAGTTAGTAATGATTATCCTTAAGT 86602
QY 270 ----- 270
Db 86601 TCATGCTTCTCAAACTAATTACATAATACATGTGTCAACATCAAATAGGCATATGCT 86542
QY 270 ----- 270
Db 86541 GTTACCAAGTGGTTCTCTATTTTCTCCTCTACAACAGACTAAATGGACAGTATCGGTT 86482
QY 270 ----- 270
Db 86481 GGCTGAGAGCATATATGCTGTTATTTTAAGTTCATCGAACTTCGCTGAAGCTGGTTG 86422
QY 270 ----- 270
Db 86421 TCCCGACGTGCACTATGCTTCTATTAAAGCTGGGCCCTTATATTTCTGTATTTG 86362
QY 270 ----- 270
Db 86361 TTTCTGGACACACCATTCATATTATATGCTGCTTATTCATAGTTACTATTATTATG 86302
QY 270 ----- 270
Db 86301 CCAATGTTTATACCAATGTCCAAATAAAATCATGTATCTACGTGAAAAATAAGTTGC 86242
QY 271 ----- 281
Db 86241 ATAAATCATTCAAATTATCAAAATTCGCGCAGAGAAAATGGCTCTTCTGTGAAGGTTTGA 86182
QY 281 ----- 301
Db 86181 GGTGGAGCTGACACTTCCATTGAAGAGTACATCATCGGACCGGCTATAGAGCCAGCTGA 86122
QY 301 ----- 321
Db 86121 TGAAGCATGTGTGATGGCGATAAGAACCAATAACACTGTACGGGCTGAAGAGGGCCA 86062
QY 321 ----- 341
Db 321 nSerTrpIleAlaArgProSerIysGlyProIleMetLeuGlySerValLeuSerIleuAl 341

Db 86061 ATCATGATGTCTGCACCTTCCAAAGGAGCAGCATTTCTTGAAGTGTCTTCTTAC 86002
QY 341 ----- 361
Db 86001 ATCTGTCATGGCAGCATGTGTGAACAGAGCGTCCACTTATGATCTTATAGTCACGCT 85942
QY 361 ----- 381
Db 85941 TTTGGCAGTGTTCATGAGAACATGCTCTCATGCTGAGAAAGCATGGCGAGTACATTGTT 85882
QY 381 ----- 401
Db 85881 eProAsnPheGlySerMetPheSerValThrAspGlnHisAlaIysAsnGlnGlnTrpAs 401
Db 85821 TCCCACTTGTGTAATGTTAGTGTGACAGATCAGACCCCAAGGTTGATCAATGGGA 85822
QY 401 ----- 421
Db 85821 PGLUGLAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyAspTyr 421
QY 421 ----- 441
Db 85761 TGAAGACAATGTCCACAGCCCGTTGCTGTCCGACAGACCAAGCGGCAGAGGGAAGGA 85702
QY 441 ----- 461
Db 85701 PLevalHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuGlu 461
QY 461 ----- 481
Db 85642 CATGCACACCACTCTCATCGTGAAGTGTCTGAGTATGAGAAAGAACCTCTTGA 85642
QY 481 ----- 501
Db 85641 YGLUGLYAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpIly 481
QY 501 ----- 521
Db 85582 eTrpSerGlnIlyGlnGlyGlnAsnGlyArgGlyGlyGlyGlyPheIysArgValTyrIle 501
Db 85581 ATGTCAGAGCGCAGAGAGCGAGGATGTGTAGAGAGAGAGATTAAAGATCTACTT 85522
QY 521 ----- 541
Db 85521 uHisGlnGlnGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 521
QY 541 ----- 561
Db 85402 GCACCAAGAGGAATGCGCAGGATCAAGAAAGGGGCTCAGTTATTCTACTTCTGTGAGG 85462
QY 561 ----- 581
Db 85461 YAspValPheGlnGlySerGlnPheValHisAlaAlaLeuValSerGlnSerAlaLe 541
QY 581 ----- 591
Db 85401 uPheSerIysGlyLeuAlaGluProArgMetSerAspAlaIleMetValHisProSerG 561
QY 591 ----- 591
Db 85281 GTTAGTCGGTGTGAATTCAATCCTTCAACAGTAGAGATGATACCAATTTCATAATT 85222
QY 591 ----- 591
Db 85221 CTTTGTGACTGTCCAACATTGACGGGTGTTAAGATTGTGTACTATTAGAATTCACT 85162
QY 592 ----- 603
Db 85161 AATTCCTTTATGTTCTTCTCTGTGACAGTTGCTGGAATAAATGGGTTCTCTATTACA 85102
QY 603 ----- 623
Db 85101 hProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerIysPheGlyLeuSerS 623
QY 623 ----- 643
Db 85041 eAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleG 643
QY 643 ----- 654
Db 643 lYPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654

Db 84981 GTTTAGCCATGAGACTTATGACATCTCTGGAAGAAGTACTAATTGTTCCCTTGCTT 84922
QY 654 ----- 654
Db 84921 GCATCATCTTAGTACTACCAATATATACCAACAAGAAATGCACCCCAAGTTAGTATTAATT 84862
QY 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleA 667
Db 84861 ACATCATACATTGTTCCCTTTTGTAGGTTTCTGCTTCTGGGACCAATTCCAGCTTGATAG 84802
QY 667 IAserLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 687
Db 84801 CATCTCTAGTGTCTGTGGTGTGTCCCAATGTTATCGACTGGGTACAGTGGCCACGCCG 84742
QY 687 euleuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProI 707
Db 84741 CACTCTCCACAATCAGCGGTCATCACTACTCTGCTGCTTCCTCATGGGATTCGGTCCGA 84682
QY 707 IeProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaI 727
Db 84681 TCCCAACATTCTGTGTGGAGATCTTCCCACTAGGGTCCGGCGCATCTGCATTGCCA 84622
QY 727 IeCysAlaPheThrPheThrIleGlyAspIleIleValThrTyrSerLeuProValMetL 747
Db 84621 TCTGCGCCCTGACATTCTGATGTGTGATATCATAGTCACTACAGCCTCCCTGATGATGC 84562
QY 747 euaAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerP 767
Db 84561 TGAATGCCATCGCGCTAGCAGGTCTTTGGTATATACGAGATGTTTGCATGCCATGCC 84502
QY 767 heValPheValPheLeuValProGluThrIleGlyMetProLeuGluValIleThrG 787
Db 84501 TCGTGTCTGCTTCTCTCAAGGTCCCGAGACGAAGGGCATGCGGCTCGAAGTCATCACCG 84442
QY 787 IuPhePheAlaValGlyAlaIleGlyAlaIleAlaIleAla 798
Db 84441 AGTTCTTCTGCTGTGTGGTCAAGCAAAATGCAGGCT 84407
RESULT 4
HVU534445
LOCUS HVU534445 2614 bp mRNA linear PLN 14-FEB-2003
DEFINITION Hordeum vulgare mRNA for hexose transporter (stpl gene).
ACCESSION AJ534445
VERSION AJ534445.1 GI:26986185
KEYWORDS hexose transporter; stpl gene.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Hordeum.
REFERENCE
1 Weschke,W., Panitz,R., Gubatz,S., Wang,Q., Radchuk,R., Weber,H. and
AUTHORS Wobus,U.
TITLE The role of invertases and hexose transporters in controlling sugar
ratios in maternal and filial tissues of barley caryopses during
early development
JOURNAL Plant J. 33 (2), 395-411 (2003)
MEDLINE 22424051
PUBMED 12535352
REFERENCE 2 (bases 1 to 2614)
AUTHORS Radchuk,R.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene
Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3,
Gatersleben, D-06466, GERMANY
FEATURES
Source
1. .2614
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/db_xref="taxon:112509"
/tissue_type="developing caryopsis"
/clone_lib="lambda ZAP2"

gene .. /dev_stage="0-15 days after flowering"
1. .2614
/gene="stpl"
CDS 83. .2314
/gene="stpl"
/codon_start=1
/product="hexose transporter"
/protein_id="CAD58958.1"
/db_xref="GI:26986186"
/translation="MSSGAVLVAAISIGNLQGMNATIAQAVLYIKKFSLETQPL
IEGLIVMSLIGATVITTFSGAVDAVGRPLLIASSVLFVSGVLMAPNVVILL
ARLIDFGIGLAVTLVPLVYISETAPDIDRLNTLPQFSGSGMFLSCVFTMSLMP
QPDWRIMGLVLSIPSLMYFALTYFLPESPRWLVSKEAKRAVLRGRDVSSE
MALVVEGLGVGKDTHEEYIIIGPDELADGLAPDOEKLYGAEQSVMIARPVKXG
GQALGSLGLMSRHGSMVSGKSLVDPLVTLFGSVHEKMPETWMSRSTLFPNFGSM
FSVAEQOAKADWDASHRDEDDYASDHGADDIEDNLSPLISROATSVGEKEIAPH
GSIMGVYESSSMQGGDAVSMGIGGQWLAKWTEREGADGQKEGFQRIYLHEBGVS
GDRGSLTSMGDDIPGGEYIQAAALVSQPALYSKDLIEQLAGPAMVHPSEAVAKG
TKWAELEPEGVKHALFVGIQLIQFAGINGVLXYTPQILEQAGVGLISNLGSSS
SASILISALTLLMLPSIGIAMRLMDSGRFLSTIPVLVLAAILLVNLVDVGT
MVHAALSTISVIVYFCFFVMGFGPIPNILCAEIPFVSRCIAICALTFWIGDIIVT
YLPVMLNAIGLAGVFGIYAVVCMIAFVFMKVPETKMPLEVITFESVGAQKQKE
ATD"
BASE COUNT 563 a 586 c 755 g 708 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 4.24e-178 Length: 2614
Score: 2808.00 Matches: 543
Percent Similarity: 84.39% Conservative: 84
Best Local Similarity: 73.08% Mismatches: 104
Query Match: 68.91% Indels: 12
DB: 8 Gaps: 6
US-10-051-909-32 (1-800) x HVU534445 (1-2614)
QY 57 SerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGluGlyTrpAsp 76
Db 89 TCGGGCGCGTGTGCTGCCCATAGCGCGCTTCATCGGAACCTGCTGCGGGGTGGAT 148
QY 77 AsnAlaThrIleAlaAlaAlaValLeuTyrIleIleIleGlyIlePheGluLeuGluAsnGlu 96
Db 149 AATGCGACCATTCGACGGTGCCTCTGTACTAAAGAAGAGTTACAGCTCGAGACCCAG 208
QY 97 ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThr 116
Db 209 CCCTTGATCGAGGGCTCATCTGTCCTCATCGAGCGACGGTTATACAGACG 268
QY 117 PheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIle 136
Db 269 TTCTCCGGGCGGTGCTGACGCTGTGTAGCGCGCCCTGCTGATCGCTGCTGTC 328
QY 137 LeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuLeu 156
Db 329 CTCTACTTTGTCAGTGGCTGTGATGCTGTGGCGCCCAACGTCTATGTGTCTCTTG 388
QY 157 AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle 176
Db 389 GCCAGGCTCATCGACGGGTTCCGTATCCGTTGGCAGTCACCTTGTCTGTACATT 448
QY 177 SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGluPheSerGly 196
Db 449 TCGGAGACCGCCCGCACTGACATTAGAGGGCTGCTGAACACGCTGCCGAGTTCAGTGGG 508
QY 197 SerGlyIleMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerPro 216
Db 509 TCAGGAGGAGTTCCTTCTTACTGATGTTTACCATTGCTCCCTCATGCCGACGACCT 568
QY 217 AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThr 236
Db 569 GACTGGAGAATCATGCTTGGGGTTTGTGCATCCCGCTTATGATATTTCATTGACT 628
QY 237 IlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMetAlaGluAla 256

```
Db      629  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      257  LysIysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyLeuSerLeu
Db      689  AAGGAGTATTGCAAGACTGGCGGAGAGGAGATGCTCAGAGAGATGGCCCTTCTT
QY      277  LeuGluGlyLeuGluValGlyAspThrSerIleGluTyrIleIleGlyProAla
Db      749  GTTGAAGGATTGGGTGTTGGAAGAAGACACACATTGTGAGAAATACATAATTGACCTGAT
QY      297  ThrGluAlaAlaAspAspLeuValThrAspGlyAspIlySerGluGlnIleThrLeuTyrGly
Db      809  GATGAGCTGGCTGATGACGGGCTGGCTCCA--GATCAAGAGAACTGAAACTCTATGGA
QY      317  ProGluGluGlyGlnSerTyrIleAlaArgProSerLys-----GlyProIleMetLeu
Db      866  GCTGAAGAGGGGTATCTTGATCGCCGCTCTGTAGGSGTGGCGGCCCAAGTGCACCTT
QY      335  GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu
Db      926  GGAAGCGCTTTGGCCCTCATGTCTCGTCATCGGAGTATGTTAGTCAGGGTAAATCTCTT
QY      355  MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly
Db      986  GTTGACCCCTTGTCACTCTCTCGGAAGTCCATGAGAGATGCTGAGGTAATGGGG
QY      375  SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis
Db      1046  AGCATGCGGAGACACTGTGTTCTTAACCTTGGCAGCATGTTTAGTGCGCTGAACAGCAG
QY      395  AlaLysAsnGluGlnTyrPaspGluAsnLeuHisArgAspAspGluGluTyrAlaSer
Db      1106  CAGGCTAAAGCTGACTGGGATGCTGAGAGT--CATAGGATGATGAAGATTATGATCG
QY      415  AspGlyAlaGlyGlyAspTyrGlyAspAsnLeuHisSerProLeuLeuSerArgGlnAla
Db      1163  GATCATGTGCTGATGACATGAGATTAACCTCAATAGCCGCTTATTCTCGTCAAGCG
QY      435  ThrGlyAlaGluGlyLysAspIleVal-----HisIleGlyHisArgGlySerAlaLeu
Db      1223  ACAAGCGTGAGGGTAAGAGATGCTGCACCTCATGGAAGCATAAATGGTGTGTGGA
QY      453  SerMetArgArgGlnThrLeuLeuGlyGlyLysAspGlyValSerSerThrAspIle
Db      1283  AGTAGTAGCATGACG-----GGAGGAGACGACGTAAGCAGCATGGGCAATT
QY      473  GlyGlyGlyTyrGlnLeuAlaTyrLysTyrSerGluLysGluGlyLysAsnGlyArgLys
Db      1328  GGTGGGGGGTGGCAGTTAGCTTGGAAGTGGAAGTGAAGAGAGAAAGTGCAGATGGGCAAAA
QY      493  GluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGly
Db      1388  GAAGGTGGCTTCAGCGGTATTACTTCATGAGAGAGGTGTCTCAGGTGATCGAGAGGC
QY      513  SerIleValSerLeuProGlyGlyLysAspValPheGluGlySerGluPheValHisAla
Db      1448  TCTATATTGCTATGCCA---GGAGGTGATATTCCCTCGGTGGTGAGTATATCCAGCA
QY      533  AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSer
Db      1505  GCGGCTTAGTGAGCCCAACCTGCTTTACTCGAAGACCTGATAGAGCAGCAGCTTGCT
QY      553  AspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeu
Db      1565  GGTCCAGCCATGTGATCATCCAGGAGGACAGTGGCCAAAGGTAACAAGTGGGACAACTA
QY      573  PheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGln
Db      1625  TTGAACTTGAGTGAAGCATGCACTGTTGTTGTCATGGAATTACAGATCCTGCAAG
QY      593  PheAlaGlyTyrLeuAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyVal
```

```
Db      1685  TTTGGGGTATCAAGGAGTCCCTGATTATACAGCCCTCAGATACTTGAGCAAGCAGTGT
QY      613  AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSer
Db      1745  GGGATTCTTATCAACATTGGACTAAGCTCTTCTCTGCATCTATTATTAGTGCC
QY      633  LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer
Db      1805  TTGACAACTTGCTGATGCTTCCACAGCATTGGCATGCGCATGAGACTCATGATATGTCA
QY      653  GlyArgArgPheLeuLeuGluGlyThrIleProIleLeuIleAlaSerLeuValIleLeu
Db      1865  GGAAGAAGTTTCTTCTCCTCTCAACAATCCCTGCTTGATAGTGACACTGACTATCTTG
QY      673  ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer
Db      1925  GTTTGTGAATGTTCTGGATGTGGAAACCATGGTGCACGCCGCTCTCAACGATCAGT
QY      693  ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys
Db      1985  GTCATAGTCTACTTCTGCTTCTTCTGTCATGGGTTTGGACCTATCCCGAATATCTCTGC
QY      713  AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaIleThrPhe
Db      2045  GCGGAGATTTCCCACTCCCTCCGTCGGGCATGTCATTGCCATCTGCGCACTAACCTTC
QY      733  TyrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu
Db      2105  TGGATTGGTGAATCATGCTGATACATACACTCTTCCGTGATGCTCAACGCCATCGGACTC
QY      753  AlaGlyValPheSerIleTyrAlaValAlaValCysLeuIleSerPheValPheValPheLeu
Db      2165  GCTGAGTTTTCGGAATATATGCTGCTGTTGTCATGATAGCCTTGTATTGCTCATCATG
QY      773  LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly
Db      2225  AAGTCCCTGAGACAAGGCGATGCCCTGAGGTGCATCACCGAGTTCTTCTCGGTGGG
QY      793  AlaLysGln 795
Db      2285  GCGAAGCAG 2293

RESULT 5
AR208568          2601 bp      DNA      linear      PAT 20-JUN-2002
LOCUS             AR208568
DEFINITION        Sequence 7 from patent US 6383776.
ACCESSION         AR208568
VERSION           AR208568.1 GI:21509755
KEYWORDS
SOURCE            Unknown.
ORGANISM          Unknown.
REFERENCE         1 (bases 1 to 2601)
AUTHORS          Allen,S.M., Hiltz,W.D., Kinney,A.J. and Tingey,S.V.
TITLE            Plant sugar transport proteins
JOURNAL          Patent: US 6383776-A 7 07-MAY-2002;
FEATURES          Location/Qualifiers
source            1..2601
BASE COUNT        648 a      543 c      644 g      766 t
ORIGIN
Alignment Scores:
Pred. No.:        3.74e-169      Length:      2601
Score:            2674.00         Matches:      522
Percent Similarity: 82.20%         Conservative: 92
Best Local Similarity: 69.88%         Mismatches:  115
Query Match:      65.62%         Indels:      18
DB:               6               Gaps:        8
US-10-051-909-32 (1-800) x AR208568 (1-2601)
QY      56  MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
```


Db	175	ATGAAAGTGCCGTCCTTGTGCTATTGTCGCGCTTCATTGCTAATTTCCTCCAGAGATGG	234
Qy	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleYsLYSGluPheGlnLeuGlnAsn	95
Db	235	GATATGCTACCATCGCGGGCTAATGTTACATTAGAAGAACCTTGCTTTGGGACA	294
Qy	96	GluproThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	295	-----ACTATGGAAGGCTTGTGTGGGCATGTCCCTGATTTGAGCAACGGTAATCAC	348
Qy	116	ThrpheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	349	ACATGCTCTGCTCCTATACGGGATTTGGCTCGGTCCGCCACCATGATGATATCTCATCT	408
Qy	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155
Db	409	GTGCTCATTTCTTGGTGGTGTGGTATGCTGTGGTCCCAATGTGATGTGTGTGC	468
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	469	TTGGCGAGGCTACTTGATGGATTTGGGATTTGGCTTGTGACTCTGTGCCGGCTTAT	528
Qy	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db	529	ATATCTGAACCGCGCCGCTCTGAATTAAGGGGGCTGTGAATAACGCTTCCTCAGTTCA	588
Qy	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Db	589	GGCTCTGAGGAATGTTTGTCTGACTGATAGTTTGGCATGTCAATTGAGTCCCGG	648
Qy	216	ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db	649	CCTAGCTGAGGCTCATGCTTGGGGTCTGTCTATTCCTTCTCTGATTTTGCAATTG	708
Qy	236	ThrIlePheTyrLeuProGluSerProArgTyrPheValSerLYSGlyArgMetAlaGlu	255
Db	709	ACCATTTTTCTTGTGCCGAGTCTCTCGGTGGCTGTCAGCAAGAAGATGCTCGAG	768
Qy	256	AlaLYsLYsValLeuGlnLYsLeuArgGlyLYsAspAspValSerGlyGluLeuSerLeu	275
Db	769	GCTAAGAGGTGCTCCAAAGATTGCGCGAAGGAGGATGTCTCAGCGAGATGGCAATTG	828
Qy	276	LeuLeuGluGlyLeuGluValGlyIleAspThrSerIleGluGluTyrIleIleGlyPro	295
Db	829	CTGCTTGAAGTCTCGGGAATGGGGGTGATACATCTATCGAAGAGTACATAATTGGCCT	888
Qy	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLYsGluGlnIleThrLeuTyr	315
Db	889	GCTGACGATGTGGCTGATGTCATGAACATGCACACAGAGAATAAATTGCATTATAT	948
Qy	316	GlyProGluGluGlyGlnSerTyrPheAlaArgProSerLYsGlyProIleMetLeuGly	335
Db	949	GGATCCCAAGCAGGCTTTCTTGTTATCAAAACCTGTCTACTGACAGAGTTCTATTGGC	1008
Qy	336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db	1009	-----CTTGGCTACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG	1056
Qy	356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Db	1057	GATCCTCTGTGACACTGTTGGTAGCATTCATGAGAAGCTCCCGAGACAGAGACAAGA	1116
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1117	GGAAGCATGCGAAGCACTGTGTTCCAAATTTTGAAGCATGTTCAAGCACTGCTGAGCCG	1176
Qy	394	HisAlaLYsAsnGluGlnTyrAspGluGluValAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1177	CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGACTACATG	1236
Qy	414	SerAspGlyAlaGlyIleAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433

Db	1237	TCAGATGCAACCCCGTGGGGACTCCGATGATTAATTGGACAGCTCCTTAAATTCACGCCAA	1296
Qy	434	AlaThrGlyAlaGluGlyLysAspLeuAlaHisGlyHisArgGlySerAlaLeu--	452
Db	1297	ACAACAAGCCTTGA---AAAGACTTACCCTCCTCCTTCCCATGGCAGTATCCTTGGC	1353
Qy	453	SerMetArgArgGlnThrLeuLeuGlyGly---GlyAspGlyValSerSerThrAsp	471
Db	1354	AGCATGAGCCGTACAGTACTCATGCAAGGGGTGAGTGAAGGTGGTAGTACAGT	1413
Qy	472	IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlyGlyLysGlyArg	491
Db	1414	ATTGGTGGCTGGCACTGGCATGGAATGACTGATAAA--GGTAGATGAAAA	1470
Qy	492	LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArg	511
Db	1471	CAACAAGAGGGTTAAAGATTTATTTACATGAGAGGAGTTCTGCATCTGCTCGT	1530
Qy	512	GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis	531
Db	1531	GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTGTCAG	1572
Qy	532	AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet	551
Db	1573	GCTGCTGCCTTGGTAAAGCAACCGCTCTTACTCCAAGAGCTTATTGATGACACCA	1632
Qy	552	SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp	571
Db	1633	GTTGGGCGCTGCATGTTCAACCATCTGAGACAGCTTCAAGGGCCAAGTTGMAAGCT	1692
Qy	572	LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln	591
Db	1693	CTTCTTGAAACCAGGGGTTAAGCATGCATTGGTTGGAGTTGGAATACAATACTTCAG	1752
Qy	592	GlnPheAlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGluAlaGly	611
Db	1753	CAGTTTTCAGGATTAATGGGGTCTATATTACACACTCAATCCTTGAAGAGCGCGT	1812
Qy	612	ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer	631
Db	1813	GTTGAAGTCTCTTTCAGATATAGGCATTGGCTGACAGTGGCATCATCTCTTATCAGT	1872
Qy	632	SerLeuThrThrLeuLeuMetIleuProCysIleGlyPheAlaMetLeuLeuMetAspLeu	651
Db	1873	GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGTT	1932
Qy	652	SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle	671
Db	1933	TCAGGCAGAGGACAGTTGCTACTTACTACAATCCCCGTGATGTGTCACTCATTAATT	1992
Qy	672	IleuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal	691
Db	1993	TTGGTCAATTGGAAGCCTGTAAATTTTGGCAATGTGGCCATGCAGCAATCTCAACAGTA	2052
Qy	692	SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu	711
Db	2053	TGCGTTGTGTTATTTCTGCTGCTTGTGTATGGGTATGCAACCAATTCCAAACATCCTT	2112
Qy	712	CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr	731
Db	2113	TGCTCAGAGATTTCCCACTAAGGTGCGTGGCCTTGCAATTGCTATGTGCAATTAGTG	2172
Qy	732	PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly	751
Db	2173	TTCTGGATTGGAGACATCATCATCACTACTCGCTGCTGTGATGCTCGCTCTTTAGGA	2232
Qy	752	LeuAlaGlyValPheSerIleTyrAlaValAlaCysIleIleSerPheValPheValPhe	771
Db	2233	CTTGGTGTGTATTCCGCATTTAACGCAAGTTGTTGTTCATCTCGTGATATTGTGTTT	2292
Qy	772	LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal	791
Db	2293	TTGAAGGTTCCAGAAACAAGGGCATGCCCTTGAAGTCACTCTGAATTCCTTTCTGTT	2352

QY 792 G1YAlAlYsglAlAlAlAlAl 798
|||||
Db 2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 6
AY094465

LOCUS AY094465 2570 bp mRNA linear PLN 05-MAY-2002

DEFINITION Arabidopsis thaliana AT4g35300/F23E12_140 mRNA, complete cds.

ACCESSION AY094465

VERSION AY094465.1 GI:20453188

KEYWORDS FLI cDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 2570)
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

AUTHORS Arabidopsis cDNA clones unpublished

TITLE 2. (bases 1 to 2570)

JOURNAL Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

REFERENCE Arabidopsis cDNA clones unpublished

AUTHORS

TITLE

JOURNAL

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

FEATURES

Source

1. 2570

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="RAFL09-48-J19(R21968)"

/note="ecotype: Columbia"

209. 2428

/note="putative sugar transporter protein"

/codon_start=1

/product="AT4g35300/F23E12_140"

/protein_id="AAM19835.1"

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

location/Qualifiers

/db_xref="GI:20453189"

/translation="MSGAVLVAlAAVGNLLQGWDNATIAGAVLYIKKEFNLESNPV
EGLIVAMSLIGATLITTCGSGVADNLRPMLILSSILYFVGSIVMLSPNVYLLLG
RLDDFGVGLVTVLPVIYISETAPPEIRGLANTLPQFTGSGGMLSYCMVFGMSLMP
PSWRLMLGLVLPISLVFELFVFLPESPRVLVSGRMLBAKRVLQRLGRBDSGEM
ALLVEGLIGGETTIEYIIGPADEVTDHDIAVDKQIKLYGAEGLSWARPVKGG
STMSVLSRHGSTMRSRQSLIDPLVTLFGSVHEKMPDTGSMRSALPFGSMFSVGN
QPRHEDWDEENLVGEEDYPSDHGDSEDDLHSPILSRQITSMRMDPHTAGTLSTF
RHGSQVQAGQEGAGSMGIGCGNQVAKMTERDESGQKEGFRILYLJEGFPGSR
GSIVSLPGDGTGEADFWQASALVSGPALYSKDLKEHTIGPAVHPSEETGSIWHD
LHDPGVKRALVVGVLQIILQFSGINGVLVYTPQILEQAGVGLISNMGISSSASL
ISALTTFVMLPAIIVAMRLMDLSGRFLTLTPIPLIASLVVISNLVHMNSIVHAV
LSTSVVLVYCFPFVMGFGPAPNLCSEIPPTRVRGICIAICALTFWICDIIVTYSLPV
LLKSGLVGFQMYAIVVCISWVFVFKVPETKGMPLVITFSPVGARQAEAAKNE"

BASE COUNT 629 a 500 c 646 g 795 t

ORIGIN

Alignment Scores:

Pred. No.: 5.02e-168 Length: 2570

Score: 2657.00 Matches: 530

Percent Similarity: 79.87% Conservative: 97

Best Local Similarity: 67.52% Mismatches: 133

Query Match: 65.20% Indels: 25

DB: 8 Gaps: 10

US-10-051-909-32 (1-800) x AY094465 (1-2570)

QY 29 LeuAlaLeuProGlyProLeuProProAlaSerCySerSerGlnGlu----- 44

Db 91 CTTCTTCTGCGCT---CCAATTCGAAGTGATTTGTTTCACTCACTCATCGTCTTGATC 147

QY 45 -----ProvalThrSerAspAspIleLeuGluAspIys----- 55

Db 148 AGTGAAGCTGTTTTCGGAATTCCTTCAATTTCTTGATAAAGATAGACGGAAT 207

QY 56 -MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTr 75

Db 208 TATGAGTGAGCTGCTTGTGCTATGCTGCTGCTGCGACACTGTTACAAAGATG 267

QY 75 PAspAsnAlaThrIleAlaAlaAlaValLeuTrIleIleIleIleIleIleIleIle 95

Db 268 GGATAACGCACTATTGACAGAGCTGTTGATACATAAAAAAGAGCTTAATTGGAGAG 327

QY 95 ngluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValTh 115

Db 328 TAATCATCATGTGAGAGGCTTAATGTGGCAATGTCACTTATTGGTACTCTGATTAC 387

QY 115 rThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSe 135

Db 388 AACATGCTCTGGAGGGGTAGCTGATGGCTGCTGCGCCGTCATGCTAATATGCTC 447

QY 135 rIleLeuTrPhePheSerGlyLeuIleMetLeuTrPheSerProAsnValTrValLeu 155

Db 448 AATTCCTACTTGTGCTTCTCTAGTAATGCTATGCTCCGAATGTTATGTGTGCT 507

QY 155 uLeuAlaArgPheValaAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTr 175

Db 508 CTTAGGAAGGTTGTATGATGATTTGGGTTGCTTGTGTGTCACACTTCTCTATTTA 567

QY 175 rIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSe 195

Db 568 TATATCTGAGACTGCACACCTGAGATTAGGGACTGTTGAATACGTCACAGTTCAC 627

QY 195 rGlySerGlyGlyMetPheLeuSerTrCyMetValPheGlyMetSerLeuSerProSe 215

Db 628 TGGCTCTGAGGATGTTCTTATCTTACTGATGTTTTCGAATGTCGTGATGCCATC 687

QY 215 rProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLe 235

Db 688 ACCTAGCTGAGATGTATGCTGTGTCCTTTTCATCCCTTCTGCTTTCTTCTCCT 747

QY 235 uThrIlePheTrIleuProGluSerProArgTrPleuValSerIleGlyArgMetAlaGl 255

Db 748 CACGGTCTTCTTCTTGCCCGAGTCCCGCAAGGTGGCTCGTGAGCAAGGTCGAATGCTTGA 807
Qy 255 uAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLe 275
Db 808 AGCAAGCGGGTCTTCAGAGACTGGGTGGTCGGAAGATGTCGTGGATGCGCTTT 867
Qy 275 uLeuLeuGluGlyLeuGluValGlyLysPheThrSerIleGluGluTyrIleIleGlyP 295
Db 868 GTTGGTTGAGGGTCTTGAATGGAGGTGAACAACACATAGAGATATATATATGTTGCC 927
Qy 295 cAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTy 315
Db 928 CGCGGATGAAGTACTGATGATCATATATAGCTGTGATAGGATCAATTAAGTTATTA 987
Qy 315 rGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuG 335
Db 988 TGGTGCAAGAAAGGGCTGAGTTGGGTGCTAGGCCAGTCAAGGA-----GG 1035
Qy 335 ySerValLeuSerLeuAlaSerArgHisGlySer--MetValAsnGlnSerValProLe 354
Db 1036 AAGCATATAGATGTTTGTCTCGCATGGAAGTACAATGAGCAGAGCAAGGCTCAT 1095
Qy 354 uMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 374
Db 1096 GATGATCCTCTTGTACACGTGGGAGCGTTCACGAGAAGATGCCGAGACT--GG 1152
Qy 374 ySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnH 394
Db 1153 AAGCATGAGGAGTGCCCTGTCCACATTTTGGAGATGTTCAGTGTGGAGGGAATCA 1212
Qy 394 sAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSe 414
Db 1213 ACCAAGACATGAAGATTGGGATGAGAAGAAATCTTGTGGAGAGGTGAGATTATCCATC 1272
Qy 414 rAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAl 434
Db 1273 CGAC---CATGAGATGATTCGAGATGATCTTCATTCTCCGTGATCTCACGCTCAAC 1329
Qy 434 aThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMe 454
Db 1330 GACAAGCATGGAG--AAAGACATGCTCACAACGTCTCATGGAACCTTTCTACCTTCA 1386
Qy 454 tArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGly 474
Db 1387 ACATGGAAGTCAAGTGCAGGGAGCTCAAGCGGGAAGGAGCGGGTATATGGGATTTGAGG 1446
Qy 474 yGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyLysAsnGlyArgLysGlu 494
Db 1447 TGGATGGCAAGTGCATGGAATGACGAAAGAGAAGATGAATCGGACAGAAAGAGG 1506
Qy 494 yGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerI 514
Db 1507 TGGGTTTAAACGATATACTTGCATCAAGAAGTTTCCAGGATCTCGACGTGCTCAT 1566
Qy 514 eValSerLeuProGlyGlyLysAspValPheGluGlySerGluPheValHisAlaAla 534
Db 1567 TGTTCATTGCTGCTGGTGTGATGAACCGGTGAG--GCAGATTGTGACAGCGTCTGC 1623
Qy 534 aLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAl 554
Db 1624 TTTGGTTAGCCCAACCAAGCTCTTATTTCCAAGACCTTCTCAAGAACATACAAATGCTCC 1683
Qy 554 aAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheG 574
Db 1684 TGTATGATCATCCATCCGAA--ACAACCTAAAGGTCATATTTGGCATGATCTTCATGA 1740
Qy 574 uProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAl 594
Db 1741 TCCCTGAGTCAGCGTGCATTAGTCGTAGAGATTGGAATTCAATACTTCAGCAGTTCTC 1800
Qy 594 aGlyIleAsnGlyValLeuTyrTrpProGlnIleLeuGluGlnAlaGlyValAlaVal 614
Db 1801 AGGCATCAACGAGTCTTACTACACACCGCAAAATCCTTGAGCAGGCGGTGTCCGGAT 1860

Qy 614 lIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuTh 634
Db 1861 CCTACTATCGAACATGGGGATTAAGTTCTTCTCCCTCAGCATCCTTACTTATAAGTGAATTGAC 1920
Qy 634 rThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyAr 654
Db 1921 AACCTTGTGATGTACTCGAATAGCTGTTGCAATGAGGCTCATGATCTTTCGTGCG 1980
Qy 654 gArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVa 674
Db 1981 AAGGACCTTGCTTCCACCAAGATACCAATCCGTGATAGCATCTTATGTTTGAATAAT 2040
Qy 674 lSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIl 694
Db 2041 CTCAATCTGTTTCACATGAACAGCATTTGTCACGCGGCTTATCAACCGTAAGCTTGT 2100
Qy 694 eValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGl 714
Db 2101 GCTTACTTCTGCTTCTTCGTGATGGGTTTCGGTCTCTCCAAACATCCTTCTTCAGA 2160
Qy 714 uIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIl 734
Db 2161 GATTTTCCAACTCGAGTCGCGGAATCTGCATCGCATCTGCGCACTACACTTCTGGAT 2220
Qy 734 eGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGl 754
Db 2221 CTGTGACATAATCGTCACTTACAGTCTCCCGTGTCTCAATCCATTGGACTGACTGG 2280
Qy 754 yValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVa 774
Db 2281 TGTGTTTGAATGATGACGCAATCGTATGTGCATTTCAATGGGCTTGTGTTCAATAAGT 2340
Qy 774 lProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 794
Db 2341 CCCGGAACATAAAGGATGCCACTTGAAGTCATCAGAGTTCTTTCTGTGGAGCTAG 2400
Qy 794 sGlnAlaAlaAla 798
Db 2401 ACAAGCTGAAGCT 2413
RESULT 7
AX506620 2190 bp DNA linear PAT 27-SEP-2002
LOCUS AX506620
DEFINITION Sequence 1315 from Patent WO0216655.
ACCESSION AX506620
VERSION AX506620.1 GI:23387857
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1315 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source location/Qualifiers
1..2190
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
BASE COUNT 529 a 439 c 572 g 650 t
ORIGIN
Alignment Scores: 2.99e-163 Length: 2190
Score: 2584.00 Matches: 512
Percent Similarity: 80.78% Conservative: 89
Best Local Similarity: 68.82% Mismatches: 123

Query Match: 63.41% Indels: 20
DB: 6 Gaps: 8
US-10-051-909-32 (1-800) x AX506620 (1-2190)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
DB 1 ATGAGTGGAGCTGTGCTTGTGCTTAATTGCTGCTGCTGTTGGCAACTTGTTACAAAGATGG 60
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
DB 61 GATAACGCAACTATTGACAGAGCTGTGTGTACATAAAAAAGAGTTTAATTGGAGAGT 120
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
DB 121 AATCATCAGTGAAGGTCTAATTGCGCATGTCACTTATGCTGCTACTGATTACA 180
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
DB 181 ACAATGCTCGAGGGGTAGCTGATGGCTTGTCGCGCTCCCATGCTAATATTGCTCTCA 240
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTropSerProAsnValTyrValLeuLeu 155
DB 241 ATTCTTACTTTGTTGTTCTCTAGTAATGCTATGCTCTCCGAATGTTATGTTGCTCTC 300
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
DB 301 TTAGGAAGGTGTTAGATGATTTGGGGTGTGCTGTGTGCACACTGTGTTCTTATTAT 360
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
DB 361 ATAATCTGAGACTGCACCACCTGAGATTAGGGGACTATTGAATACGCTACCGCAGTTCACT 420
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
DB 421 GGCTCTGAGGGATGTTCTTATCTTACTGATGTTTCGGAATGTCGTTGATGCCATCA 480
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
DB 481 CCTAGCTGAGATTGATGCTGTGCTCTTTCATCCCTTCCTTCTTCTCTC 540
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
DB 541 ACGGTCTTCTTCTTCCCGAGTCCCAAGGTGGCTCGTAGCAAAAGTCGAATGCTTGAA 600
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
DB 601 GCAAAAGCGGGTCTTCAGAGACTCGGTGTCGCGAAGATGTCTGTGAGATGGCTTGG 660
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
DB 661 TTGGTTGAGGGTCTTGAATTGAGGTGAACAACACATAGAGGAATATATAATTGTCCTCC 720
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
DB 721 GCGGATGAAGTTACTGATGATCATGATATAGCTGTGATTAAGATCAATTAAGTTATAT 780
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
DB 781 GGTGCAGAAAGAGGCTGAGTTGGGTTGCTTAGGCCAGTCAAAGAA-----GGA 828
QY 336 SerValLeuSerLeuAlaSerArgHisGlySer--MetValAsnGlnSerValProLeu 354
DB 829 AGCACTATGAGTGTGTTGTCTCGCATGGAAGTACMATAGACAGAGGCAAGGCTCATTG 888
QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
DB 889 ATTGATCCTCTTGTCACACTGTTGGGAGCGTTCAAGAGAGATGCCGACACT--GGA 945
QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
DB 946 AGCATGAGAGTGCCTGTGTCACATTTTGGAGATGTTCAGTGTGGAGGGAATCAA 1005

QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
DB 1006 CCAAGACATGAAGATTGGAGTGAAGAGAAATCTTGTGGAGAAGGTGAGGATTAATCCATCC 1065
QY 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
DB 1066 GAC--CATGAGATGATTTCTGAAGATGATCTTCATCTCCGTTGATCTCAAGTCAAAACG 1122
QY 435 ThrGlyAlaGluGlyLysAspIleValHisIleGlyHisArgGlySerAlaLeuSerMet 454
DB 1123 ACAAGCATGAG--AAAGACATGCTCTCACACTGCTCATGGAACCTCTTCTACTTCAGA 1179
QY 455 ArgArgGlnThrLeuLeuGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
DB 1180 CATGGAAGTCAAGTGCAGGGAAGCTCAAGGGGAAGAGCGGGTAGTATGGGATTTGAGGT 1239
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494
DB 1240 GGATGGCAAGTGGCATGGAATGGACGGAAGAGAAATGAATCGGACAGAAAGAA-- 1296
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
DB 1297 -----GAAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329
QY 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534
DB 1330 GTTTCATGCTCGTGGTGATGGAACCGGTAG--GCAGATTTGTACAGCGTCTGCT 1386
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
DB 1387 TTGGTTAGCCAACCAAGCTCTTATTCCAAAGACCTTCTCAAGAACAATACAAATTGGTCTCT 1446
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574
DB 1447 GCTATGATCATCATCCGAA--ACAACCTAAAGGGTCAATTGGCATGATCTTCATGAT 1503
QY 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594
DB 1504 CCTGAGTCAAGCGTGCATTAGTGTGAGGATTTGGAATTCAAAATACCTTCAGCACTTCTCA 1563
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
DB 1564 GGCATCAACGAGTCTTACTACACACCGCAATCCTTGAGCAGCGGGTGTCCGGATC 1623
QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
DB 1624 CTACTATCGAACATCGGGGATTAAGTCTTCTTCACGATCCTTACTTATTAAGTCATTGACA 1683
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
DB 1684 ACCTTGTGATGTTAACCTGCAATAGCTGTGGCAATGAGGCTCATGATCTTCTGTGTCGA 1743
QY 655 ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
DB 1744 AGGACCTTGCTTTCACACGATACCAATCGATGATGATCTTATGTTTGAATATC 1803
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
DB 1804 TCAAAATCTTGTTCACATGAACAGCATGTGTCACGCGGTCTTATCAACCGTAAAGCTTGTG 1863
QY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
DB 1864 CTCTACTTCTGCTTCTTCTGATGAGGTTTCGCTCTGCTCCAAACATCTCTGTTCAGAG 1923
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
DB 1924 ATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCGCACCTTCTTGATC 1983
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
DB 1984 TGTGACATAATGTCATTACAGTCTCCCGTGTGCTCAAAATCCATTGAGACTGCTGT 2043
QY 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheLeuLysVal 774

Db	Accession	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
Db	2044 GTCGTTGGATGTACCGCAATCGTATGTTGCATTTTCATGGGTCTTTGTGTTCAATTAAGTC	2103	2584.00	80.78%	68.82%	63.41%	8
OY	775 ProGUThrLySGlyMeCProLeuGLuValIleThrGluPhePheAlaValGlyAlaLys	794					
Db	2104 CCGAAACTAAAGGCATGCCACTTGAAGTCATCACAGACTTCTTTCTGTTGGAGCTAGA	2163					
OY	795 GlnAlaAlaAla 798						
Db	2164 CAAGCTGAAGCT 2175						
RESULT 8	ATH532570	2190 bp	mRNA	linear	PLN 11-DEC-2002		
LOCUS	ATH532570						
DEFINITION	Arabidopsis thaliana mRNA for monosaccharide sensing protein 2 (mssp2 gene).						
ACCESSION	AF532570.1	GI:26800694					
VERSION							
KEYWORDS	monosaccharide sensing protein 2; mssp2 gene.						
SOURCE	Arabidopsis thaliana (thale cress)						
ORGANISM	Arabidopsis thaliana						
REFERENCE	1 Stamme, C., Tjaden, J., Trentmann, O., Emmerlich, V. and Neuhaus, E. A novel family of plant monosaccharide transporters is involved in a new type of eukaryotic sugar sensing						
AUTHORS	2 (bases 1 to 2190)						
TITLE	Neuhaus, E.						
REFERENCE	Submitted (04-DEC-2002)						
AUTHORS	Neuhaus E., Plant Physiology, University of						
TITLE	Kaiserslautern, Erwin-Schroedinger-Str. 22, 67663 Kaiserslautern,						
JOURNAL	GERMANY						
FEATURES	Location/Qualifiers						
SOURCE	1..2190						
gene	/organism="Arabidopsis thaliana"						
CDS	/mol_type="mRNA"						
	/db_xref="taxon:3702"						
	1..2190						
	/gene="mssp2"						
	1..2190						
	/gene="mssp2"						
	/codon_start=1						
	/product="monosaccharide sensing protein 2"						
	/protein_id="CAD58692.1"						
	/db_xref="GI:26800695"						
	/translation="MSGAVLVVAIAAVGNLLOGMDNATIGAVLYIKKEFNLESNP						
	EGLIVAMSLIGATLITTCGGVADWLGRRPMLTSLIFVGLVMLWSPNVYVLLLG						
	RLDGFVGVLVTLVPIYISETAPEIRGLNLPQFTSGGMFLSYCMVFGMSLMPS						
	PSWRMLGVLFIPSLVEFFLTVFFLPESPRMIVSKRMLAKRVLQRLGRREDVSGEM						
	ALLVEGLIGGETTIEEYIIGPADEVTDHDAVDKQIKLYGAEEGLSWARPVKGG						
	STMSVLSRHGSTMSRRQSLIDPLVTLFGSVHEKMDPTGSMRSLFPHFGSMFSVGN						
	QPRHEDWDEENLVGEGEDYPSDHGDDSEDLHSLPLISROTTSMEKMDPTAHGTLSTF						
	RHGSOVQAGAGGEGAGSMGIGGMQVAMKTERDESGOKEEGFPGSRGSLVSLPGAD						
	GTGEADFVQASALVSQPALYSKDLKEHTTIGPAVHPSETTKGISIWHDLHDPGVKRAL						
	VGVGGLQILQFSGINGVLVYTRQILKEAGVGLISNMGISSSASALISALTTFVML						
	PAIAVAMRLMDLSEGRRTLLLTTPILIASLVVISNLVHMSIVHAVLSTVSVLYF						
	CFEVMGFGEPAENILCEIPEPTRVRGICIAICALTFWICDIIVTYSLPVLLKISIGLAGV						
	FGMYAIVCCISWVEVFIKVPETKGMPLVETTERFSVGARQAEAAKNE"						
BASE COUNT	529 a 439 c 572 g 650 t						
ORIGIN							
Alignment Scores:	2.99e-163	length:	2190				
Pred. No.:	2584.00	Matches:	512				
Score:	80.78%	Conservative:	89				
Percent Similarity:	68.82%	Mismatches:	123				
Best Local Similarity:	63.41%	Indels:	20				
Query Match:	8	Gaps:	8				

QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr	75
Db	1	ATGAGTGGAGCTGTGCTTGTCTATTGCTGTGCTGTGGCAACTGTITCAAGAGATGG	60
QY	76	AspAsnAlaThrIleAlaIleValLeuTyrIleIleGlyGlnPheGlnLeuGlnAsn	95
Db	61	GATAACGCAACTATTGCAGAGAGCTGTGTGTACATAAAAAAGAGTTTAATTGGAGAGT	120
QY	96	GluProThrValGlnGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	121	AATCCATCAGTGGAGAGTCTAATTGTGGCCGATGTCACCTATTGTGTCTACTCTGATTACA	180
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	181	ACATGCTCTGGAGGGGTAGCTGATTTGGCTGTGCGCCGTCCTCCATGCTAATATTGCTCTCA	240
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPseProAsnValTyrValLeuLeu	155
Db	241	ATTCTCTACTTTGTGTCTCTAGTAATGCTATGCTCTCCGAATGTTTATGTGTGCTC	300
QY	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	301	TTAGGAAGGTGTGTAGTNGATTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAT	360
QY	176	IleSerGlnIleAlaProSerGlnIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db	361	ATATCTGAGACTGCACCACCTGAGATTAGGGGACTGTGAATACGCTACCGCAGTTCACT	420
QY	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Db	421	GGCTCTGAGGAGATGTTCTTACTTACTGATGTTTCGGAATGTCGTTGATGCCATCA	480
QY	216	ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db	481	CCTAGCTGAGATGATGATGCTGTGTCTTTTCACTCCCTTCCCTGTCTTTTCTTCCTC	540
QY	236	ThrIlePheTyrLeuProGlnSerProArgTyrLeuValSerIleGlyArgMetAlaGlu	255
Db	541	ACGGTCTTCTTCTTCCCGAGTCCCAAGGTGGCTGTGACCAAGGTGCAATGCTTGAA	600
QY	256	AlaIleGlyValLeuGlnIleGlyLeuArgGlyIleGlyAspAspValSerGlyGluLeuSerLeu	275
Db	601	GCAAAGCGGTTCTTCAGAGACTGCGTGTGCGCAAGATGTGTGTGTGAGATGGCTTTG	660
QY	276	LeuLeuGlnGlyLeuGlnValGlyGlyAspThrSerIleGlnGluTyrIleIleGlyPro	295
Db	661	TTGCTTGAGGCTCTTGAATTTGGAGGTGAACAACATAGAGGATATATATATGTCCTCC	720
QY	296	AlaThrGlnAlaIleAspAspLeuValThrAspGlyAspIleGlnIleThrLeuTyr	315
Db	721	GCGGATGAAGTTACTGATGATCATGATATAGCTGTGATTAAGGATCAAAATTAACTTAAT	780
QY	316	GlyProGlnGlyGlnSerTyrIleAlaArgProSerIleGlyProIleMetLeuGly	335
Db	781	GGTGCAGAGAGGCTGAGTTGGTGTGAGGCCAGTCAAAAGA-----GGA	828
QY	336	SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu	354
Db	829	AGCACTATGAGTGTGTTGTCTCGCCATGGAAGTCAATGACAGAGGCAAGGCTCATTTG	888
QY	355	MetAspProIleValThrLeuPheGlySerValHisGlnAsnMetProGlnAlaGlyGly	374
Db	889	ATTGATCCTCTTGTCACTGTTTGGAGCGCTTCAAGAGAAGATGCCGAGACACT---GGA	945
QY	375	SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis	394
Db	946	AGCATGAGGAGTGCCTTGTTCACACATTTTGGAGTATGTTTCAAGTGTGAGGGAATCAA	1005
QY	395	AlaIleAsnGlnGlnTyrPaspGlnGluAsnLeuHisArgAspAspGlnGluTyrAlaSer	414
Db	1006	CCAAGACATGAGATTTGGATGAAGAGAACTTTGTTGAGAGAAGGTGAGGATTAATCCATCC	1065

QY 415 AspglyAlaGlyGlyAspTyrGluaspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db 1066 GAC---CATGAGATGATTCGTGAAGATGATCTTCATTCCTCGTTGATCTCACGTCAAACG 1122
QY 435 ThrGlyAlaGluGlyLysAspIleValHisIsgLynHisArgGlySerAlaLeuSerMet 454
Db 1123 ACAAGCATGGAG--AAAGACATGCTTCACACTGCTCATGGAACCTCTTCTACCTTCAGA 1179
QY 455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
Db 1180 CATGGAAGTCAAGTGCAGGAGCTCAAGGGAGAGAGCGGGTAGTATGGGATTTGAGGT 1239
QY 475 GlyTyrGlnLeuAlaTrrplySerTrpSerGluLysGlyGlyLysGlyArgGlyGlyGly 494
Db 1240 GGATGGCAAGTGGCATGGAAATGGACGGAAAGAGAAGATGAAATCGGACAGAAAGAA--- 1296
QY 495 GlyPheLysArgValTyrLeuHisGlnGlyValProGlySerArgArgGlySerIle 514
Db 1297 -----GAAGGTTTCCAGGATCTCGACGTGCTCAATT 1329
QY 515 ValSerLeuProGlyGlyGlyAspValPheGlySerGluPheValHisAlaAlaAla 534
Db 1330 GTTTCATTCCTGCTGTGATGGAACCGGTGAG---GCAGATTTTGTACAAGCCTGCT 1386
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
Db 1387 TTGGTTAGCCCAACGAGCTCTTATTCCAAAGACCTTCTCAAGAACAATACATATTGGTCT 1446
QY 555 AlaMetValHisProSerGlyValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574
Db 1447 GCTATGCTACATCCAGAA---ACAACATAAAGGTCATATTGGCATGATCTTCATGAT 1503
QY 575 ProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594
Db 1504 CCTGGAGTCAAGCGTGCAATTAGTCGAGAGTTGACTTCAATACTTCAGCAGTTCTCA 1563
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAlaVal 614
Db 1564 GGCATCAACGGAGTTCTTACTACACACCGCAAACTCTTGACAGCGGGTGTGGGATC 1623
QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
Db 1624 CTACTATCGAACATGGGATTAGTCTTCTCAGCATCCTTACTTATAAGTCATTGACA 1683
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
Db 1684 ACCTTTGTGATGTTACCTGCATAGCTGTTCGAATGAGGCTCATGTGATCTTCTGCTCA 1743
QY 655 ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
Db 1744 AGGACCTTGCTTCTCACACGATACCAATCTGTATAGCATCTTATGTGTTTAGTAATC 1803
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
Db 1804 TCAAACTTGTTCACATGAACAGCATGTGTGACGCGGTCTTATCAACCGTAAGCGTTGTG 1863
QY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
Db 1864 CTCTACTTCTGCTTCTTCGTGATGGGTTTCGGTCTCTCCAAACATCCTCTGTTCAGAG 1923
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
Db 1924 ATTTTCCAACTCGAGTCCGGGAAATCTGCATCGCCATCTGCGCACTCACCTTCTGATC 1983
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
Db 1984 TGTGACATAATCGTCACTTACAGTCTCCCGGTGCTGCTCAAAATCCATTGAGCTAGCTGT 2043
QY 755 ValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLysVal 774
Db 2044 GTGTTTGAATGTACGCAATCGTATGTGATTTTCATGGGCTTTGTGTTCAATTAAGTC 2103
QY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794

Db 2104 CCGGAACATAAGGCGATGCCACTTGAAGTCATCATCAGAGTTCTTCTGTGAGCTAGA 2163
QY 795 GlnAlaAlaAla 798
Db 2164 CAACTGAAGCT 2175
RESULT 9
AC073166
LOCUS
DEFINITION
Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence,
complete sequence.
AC073166
AC073166.7 GI:12039314
HTG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharthoideae; Oryzeae; Oryza.
1 (bases 1 to 142114)
Buell,C.R., Yuan,Q., Moffat,K.S., Hill,J.N., Burr,P.C., Hsiao,J.,
Zsammann,V., Pai,G., Bowman,C.L., Fujii,C.Y., VanAken,S.B.,
Bowman,C.L., Craven,B., Utterback,T.R., Khalak,H., Feldblyum,T.V.,
Quackenbush,J., White,O., Salzberg,S.L. and Fraser,C.M.
Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence
Unpublished
2 (bases 1 to 142114)
Buell,R.
Direct Submission
Submitted (09-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 142114)
Buell,R.
Direct Submission
Submitted (05-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
4 (bases 1 to 142114)
Buell,R.
Direct Submission
Submitted (28-AUG-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
On Jan 5, 2001 this sequence version replaced gi:11968438.
Address all correspondence to:rice@tigr.org
COMMENT
BAC clone OSJNB0064P21 is from Oryza sativa chromosome 10
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Genscan and Genscan+ (Chris Burge,
http://CCR-081.mlt.edu/GENSCAN.html), GenemarkHMM (Mark Borodovsky,
http://genemark.biology.gatech.edu/Genemark/), and Genesplicier
(Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (http://www.tigr.org/cdb/tgi.shtml).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
FEATURES
source
1..142114
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"

/chromosome="10"
/map="RZ500"
/clone="OSJNB0064P21"
complement(3008..3658)
/gene="OSJNB0064P21.16"
/note="predicted by genscan+"
complement(<3008..>3658)
/gene="OSJNB0064P21.16"
complement(3008..3658)
/gene="OSJNB0064P21.16"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA046103.1"
/db_xref="GI:12039315"
/translation="MVAGDSQARLLVTLALLRLLDPAASAAPELFRHSDYRATV
ARGISVDFVWAPFESNLTRLLHEDRLAPRTPDVLVSGSLWMLHVTDAARYGDALA
SVVDAKSLRSPLVPVPPHFWLGLPLVNHMLTDKVKVMNDTILQAYDLEVEQRG
HLQRDGGPFLLDVGKLSRGGCGQCTADGMHYDGDVYDAVLHIMNALVIESQRI"
complement(3574..3596)
/rpt_family="GC_rich"
complement(4269..4297)
/rpt_family="GC_rich"
5613..5633
/rpt_family="GGAA)n"
complement(5634..6005)
/gene="OSJNB0064P21.1"
/note="identical to histone H4 (H4C13) GI:168499 (Zea
mays), (Plant Physiol Biochem 25, 235 247 (1987));EST
D23040 from this gene"
complement(<5634..>6005)
/gene="OSJNB0064P21.1"
complement(5634..5945)
/gene="OSJNB0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AA046106.1"
/db_xref="GI:12039318"
/translation="MSGRGKGKGLCKGAKRHRKVLKDNIQITKPAIRRLARRGV
KRISGLIYBTRGVLKIFLENVIRDAVTYTEHARRKVTAMDVYALKRQGRLLYFGG
G"
complement(7473..7546)
/rpt_family="CGG)n"
7556..7630
/rpt_family="GGAGAA)n"
complement(8238..8259)
/rpt_family="AT_rich"
complement(8915..8963)
/rpt_family="AT_rich"
complement(11055..11097)
/rpt_family="AT_rich"
complement(11244..11264)
/rpt_family="AT_rich"
14313..18336
/gene="OSJNB0064P21.2"
/note="identical to calcium dependent protein kinase
GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993));
EST C22538, C22540, AA750452, C99271, C22539 from this
gene"
join(14313..15004,15338..15481,15564..15716,15894..16009,
16646..16810,17112..17333,17903..18336)
/gene="OSJNB0064P21.2"
join(14323..15004,15338..15481,15564..15716,15894..16009,
16646..16810,17112..17333,17903..18025)
/gene="OSJNB0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AA046110.1"
/db_xref="GI:12039322"
/translation="MGNSCQNGTYGNNYQNSNRFQNDPRASRYVDGNDTEDCYSGSSR
ASLAGALROGLNLKSPVLGYKTFNVREL YTLGRELGGQGGKTYLCTEISTGCGYACK
TILKSNLRCVSDIEDVREIQLMHLSGOKNIYTIKDTYDEQAVHIVMELCAGELF
SKIQRGHYSERKAELIKIYGIETCHSHGVMDRLKENFLLLDADDEFSVKALD
FGLSVFFRRGQVFRREVGSPTYIAPEVLEKRYGPEADITWAGVILYVLLTGVPFWAD

TQSGIYEKVLIDGRIDFKSNRWPRISDAKDLIKMLCPYPSERLKAHEVLKHPWICDN
GVATNRALDPSVLPRLKQFSAMNRLKLSLQIIAERLSEEEIVGLREMFKAMDTKRS
VVFEGELKGLKRYSSVEKDTIINDLEAADTSTINWEEFIAAVSINKTEREKLIM
AAFTYFDKDGSGFITVDKLOKACMERNMEDTFLEMITLEVQNDNGQIDYAEFVTMMQ
SNFGLGWQYVESSLINVALREAPQVY"
complement(18675..18732)
/rpt_family="GC_rich"
complement(20605..20634)
/rpt_family="AT_rich"
22393..24891
/gene="OSJNB0064P21.6"
/note="predicted by genemark hmm"
<22393..>24891
/gene="OSJNB0064P21.6"
22393..24891
/gene="OSJNB0064P21.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA046113.1"
/db_xref="GI:12039325"
/translation="MAGSGGEAMASSGLDGEYLDRPDAIHRRAASVAIVRSGGDP
RIVDGRDPRARSSRRLSLSSWRPAGSETSVESRRYGFYTAASAAAEVGGKTRAAI
EMKPVRLSRIGKRRCGLLSVEAGVGVASMDGLRLAVTVRAEYKQDGAVQTMPA
TVRGGCADFVETLLFVRCNITYPAGAGTGPRLKLEPRFVSVVPAPAPVRLGAHVD
VSSLVDSLQKSSSEGRVRWFDTAVTLGKATGCELLKLGFGQMDAGLCLYTQAT
EKVDVVDVSPALARAHNKNSFSVARTSGPKLSASDAISPMSRAYQLIDRLSVDEH
GDPVTSILPRKLADDELSDVGLPEYEVVDKVEYTVKEVHYHAHRDYLKELDSIAEO
IEAIEALMTNGKSPSPSKVDQKOCADDEMTVEFLRKLAEVVDKGRKLKQPM
PRSESEKKAASAAAPVVPDLGPGLTAVQTRDGGFLVSMNPFDLPLEKDAAPKIAMQ
VSRPFVLSSMAATGEDVLQKMAAGADEVDKVARLGMNDLTGKTPBOVFEGIA
EAVIGGRTEGASSAARSVRLVRKLAASVDSRSESVATGIWTAADDEPTEEVIAF
SLQKLEAMAVDALMIQEMADEDAFEVAPAGDATTVFDALVSPDEMSERSGSDGRV
TVVAALQLRDPSSRRYEAVPMVAVVQSAKLLGAAGNSGGRFVRSIHHVGGVQLRCPV
GGAGCGRASNGAEROKLTAMQWALAHGPAPARAAGRARTPTTTTPSSGARQRPDVVW
SLSSRVLAGWMLKTVRNPDRVGVGATAAGSGGGN"
complement(23766..23788)
/rpt_family="GC_rich"
complement(24616..24639)
/rpt_family="GC_rich"
complement(24693..24714)
/rpt_family="GC_rich"
24853..24903
/rpt_family="CGG)n"
complement(26149..26194)
/rpt_family="AT_rich"
complement(26861..26905)
/rpt_family="AT_rich"
complement(27502..27537)
/rpt_family="GC_rich"
/rpt_family="GC_rich"
complement(27628..27663)
/rpt_family="AT_rich"
complement(28100..28132)
/rpt_family="AT_rich"
complement(29892..29929)
/rpt_family="TA)n"
29893..29928
/note="with dinucleotide (AT) repeat (35 to 42 copies)"
34449..38345
/gene="OSJNB0064P21.3"
/note="similar to sugar transporter GI:1495273
(Arabidopsis thaliana); EST AA231788, D48171, C72446,
AA749640 from this gene"
join(34449..34603,34798..34894,35286..35852,35977..36929,

Alignment Scores:
Pred. No.: 6.01e-156 Length: 142114
Score: 2514.00 Matches: 552
Percent Similarity: 53.77% Conservative: 75
Best Local Similarity: 47.34% Mismatches: 109
Query Match: 61.69% Indels: 431
DB: 8 Gaps: 9
US-10-051-909-32 (1-800) x AC073166 (1-142114)

```
QY      56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db      34816 ATGGCGGGCGCCGCTGCTGTCGCCATCGCGCCCTCCATCGGCAACTTGCTGCAGGGCTGG
QY      76 AspAsnAlaThrIleAla----- 81
Db      34876 GATTAATGCAACCATGTGACGTAAGTTGCTGTTAGAGAAAAACGAATCTTTTCTTCT
QY      81 ----- 81
Db      34936 GCCATAGTTTGTGTCCATGTCTTAATGTGACGCTAGGGCCCGTGTGTCGCAATTAT 34995
QY      81 ----- 81
Db      34996 TGTGTAATCTGTGAGATCATGAGATTGGTGAATCATCTGCAGTAAATAAGTTT 35055
QY      81 ----- 81
Db      35056 AGTTGTCTTCTCTGTTCATGATATCAAAAAAACAACTGATATTTTTCAGTA 35115
QY      81 ----- 81
Db      35116 GAATTTTGAGTCAATTATGAGTAGAATTAAACAATGTACCGATCTCATGAATGATCTGT 35175
QY      81 ----- 81
Db      35176 AATTATGGAACCTTGATACTGTCTGCATCGTAAGGTTTCAGAAATTGATGATACCGATG 35235
QY      82 -----AlaIaValle 85
Db      35236 GTTTTGACTTCTTGCTGACCCCTGTGCACTTGTGGGATGATCTTATCAGGTGCGGTACT 35295
QY      85 uTyrIleIleLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSe 105
Db      35296 GTACATCAAGAAGAATTCAACTTGACAGAGCGAGCCCTTATCGAAGCCCTGATCGTGGC 35355
QY      105 rMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerI 125
Db      35356 CATGTCTCATTTGGGGCGACGATCATCAAGACGTTCTCTGGACAGTGGCTGATTTCTT 35415
QY      125 eGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe 145
Db      35416 TGGTAGCGCGCCCATGTGATCGCGTGGCTGTCTCTACTTGTGAGTGGCTAGTGAT 35475
QY      145 tLeuTyrSerProAsnValTyrValLeuLeuValAargPheValAspGlyPheGlyI 165
Db      35476 GCTTTGGCGCGCAAAATGTATGTGTGCTCTTGGCAGGCTCATTGACGGGTTCCGGAT 35535
QY      165 eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr 185
Db      35536 CGGTTTGGCTGTCAAGCTTGATCATCTCTGAGACTGCCCCGACGACATCAG 35595
QY      185 gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCy 205
Db      35596 AGGACTGCTAAACACGCTGCCGAGTTCAGTGGGTCTGGAGGAGTTCCTTTCATACTG 35655
QY      205 sMetValPheGlyMetSerLeuSerProSerProAspTyrPargIleMetLeuGlyValLe 225
Db      35656 CATGTATTTGGCATGTCCCTCATCGCCACAGCCAGATTGGAGGATCATGTGGCGTTCT 35715
QY      225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProAr 245
Db      35716 ATCAATACCATCACTTATATACTTTCATTTGACCATCTTTTACTTGAATCGCCGAG 35775
QY      245 gTyrLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgG 265
Db      35776 GTGGCTCGTGAGCAAAAGAAATGGCTGAGGCCAAGCGTGTGTGCAAGGCTGCGTGG 35835
QY      265 yLysAspAspValSer----- 270
Db      35836 AAGACAAGATGTTTCAAGTGTGTATGTCTTACCACGATTTTCCCTGTGAGGATCTAGA 35895
```

```
QY      270 ----- 270
Db      35896 AAGTGAGAAACTATTTAATGCTATCAGTATAGTATATTATCTATGCGATGAGTTTCT 35955
QY      271 -----GlyIleuSerLeuLeuLeuGlnGlyLeuGluValGlyG 284
Db      35956 CATGTGATTAACCTGTCAGAGAAATGCGCCCTCTCGTTGAAGTCTGGGGTTGGGA 36015
QY      284 yAspThrSerIleGlnGluTyrIleIleGlyProAlaThrGluAlaAspAspLeuV 304
Db      36016 AAGACACAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTGCTGATGAAGGC 36075
QY      304 aThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGlnGlyGlnSerTyrI 324
Db      36076 TGGCTCCA--GATCCAGAGAAGATCAAACTGTATGCTCTGAAGAAAGGCTTATCGTGG 36132
QY      324 leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH 344
Db      36133 TTGCCCGTCTGTTACCGGGCAAAAGTGCACCTTGGAGTCATTAGGTCCTCATCTCTCGTC 36192
QY      344 lsgIySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS 364
Db      36193 ATGGTAGTATGTCAGTCAGGGTAAGCCCTTGTCGATCCTGTGTACCCCTTTTGAA 36252
QY      364 eValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnP 384
Db      36253 GTGTCCATGAGAAGATGCTGAGATAATGGAGACATGCGGAGACATGTGTTCTTAAC 36312
QY      384 heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTyrPaspGluGluA 404
Db      36313 TTGGCAGCATGTTTATGTGTGGCGGCAACAGCACGACCTAAAGGTGATTGGGTGCTGAGA 36372
QY      404 snLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspA 424
Db      36373 GT--CAACGGAGGGTGAAGATTATGATCAGACCATGCTGGGAGTGCATTGAAGATA 36429
QY      424 snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle--V 443
Db      36430 GCCTCCAAGCCCACTTATTCTCGTCAAGCGACAAAGCTGGAAGAAAGAGATCGCTG 36489
QY      443 aHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluG 463
Db      36490 CACCTCATGGCAGTATATAGGCTGTGGGA-----AGAAAGTAGTAGTCTCATGACAG 36543
QY      463 lylGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTyrLysTrpS 483
Db      36544 GCGGGGAGGACAGTAAGACGATGGGCATTGTGTGGGGATGGCAGTGGCTTGAATAATGA 36603
QY      483 eArgLysGlnGlyGluAsnGlyArgLysGlyGlyPheLysArgValTyrLeuHisG 503
Db      36604 CTGAGAGAGAAAGTGCAGATGGCGAAAAGAGGTGGCTTCCAAAGTATCTACTTGCATG 36663
QY      503 lngLuglyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyLysAspV 523
Db      36664 AAGAGGTTGTACAGGTGATCGAGGGCTTATACTGTCAATTGCT--GGAGGTGATG 36720
QY      523 alpHeGlnGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheS 543
Db      36721 TTCCTCTGTGTGAGTTGCTCCAGGACGTGCTTGTCAAGCCAACTGCTCTTACT 36780
QY      543 eRlySGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValA 563
Db      36781 CTAAGGAATTGATGAGCAACGCTTGCTGCGCCCTGATAGTGTCATCCATCTCAGGCAG 36840
QY      563 laAlaLysGlySerArgTyrLysAspLeuPheGluProGlyValArgArgAlaLeuLeuV 583
Db      36841 TTGCTAAAGGTCCAAAATGGGACAGACTTATTGCAACTGGAAGTGAAGCATGCTCTGTTTG 36900
QY      583 aGlyValAlGlyIleGlnIleLeuGln----- 591
Db      36901 TTGGCATAGGATACAAATCCTGCAACAGTAACCTTTTCTGATTTTATGTCATGTAC 36960
QY      591 ----- 591
```

Db 36961 GAATGCATGCTATGTTGAACCTAATTTCTAGTAGCTCCAGTATAAATTACTCTGAT 37020
QY 591 -----
Db 37021 TAATGTAAGTGTAAATCAAGCATCTTCATGCAGTGTCTCTAGGTTCTAGCGAG 37080
QY 591 -----
Db 37081 TCTCTATTAAAGCTCTTACTTACTCTACTCGAAAAACATGTTTAGCTGTACGTTT 37140
QY 591 -----
Db 37141 TCTTGATATATATTCATACAGGACAATGAGCTAACTGCACAATAATACACCTTG 37200
QY 591 -----
Db 37201 ATGATTAGAATACAATGCGATCAATGTTTGTAGTTTACACTAGTAAGTCAACCT 37260
QY 591 -----
Db 37261 TAACATCTTGCAATATCTATTAAGAAAACCATGGAGCTGGCTATGTCACATGACC 37320
QY 591 -----
Db 37321 TGATTACGGGTATTCTATCTGAAGCAGAAACAATGTGATGATCTTTATTCCTG 37380
QY 591 -----
Db 37381 AATTATTTGCTGAATATGTTCTAGTTTCGACGAATAACTTTCAGAACTTTCACAGT 37440
QY 591 -----
Db 37441 GGCATATCATGATATCTCTTTTACTTCCCTCCACTCTTCCCTATTTTCCCTGTAT 37500
QY 591 -----
Db 37501 ATTTATATCAGTCTATTCATGTTTACCCTCACTATAGCATACTGAATGTTAATGTCTAT 37560
QY 592 -----
Db 37561 CAATTTTCCATTATGTCAGATTGCTGGCAATTATGAGTTCTGTACTACACTCCACAA 37620
QY 606 TLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
Db 37621 ATTCTTGAGCAAGCTGTGTTGTTCTTCTTGGCAACATGACTTAGCTCTCTCATCT 37680
QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuMetLeuProCysIleGlyPheAla 645
Db 37681 GCATCTATTCTTATTAGCGAGCTGACCAACCTGTCTGATGCTTCCAGCATGTGATGTCT 37740
QY 646 MetLeuMetAspLeuSerGlyArgArg----- 655
Db 37741 ATGAGGCTCATGATATGCTTGAGAGAGAGGTTAGACTTCATATTTTACATTTTGTGT 37800
QY 655 -----
Db 37801 GATTACATGCATCACTTCTAAGTGAAGTATATGCTAACCATTAATTCGCGTGTCTCT 37860
QY 656 -----PheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 37861 CATTCAGGTTTCTT-CTCCTTGACAACATCCCTATCCTGATAGTACACTAGCTATCTTG 37919
QY 673 ValValSerAsnLeuIleAspLeuGlyThrIleAlaHisAlaLeuLeuSerThrValSer 692
Db 37920 ATCTGTGCAATATCTGTGATGTGGGACCATGTTTCATGCTCCTCACAGTCACT 37979
QY 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 37980 GTCATACTCTACTTCTGCTTCTTGTTCATGGGGTTCGGGCTTATCCAAACATCTCTGT 38039
QY 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732

Db 38040 GCAGAGATTTCCCGACACCCTGCTGCGATCTGCATAGCCATCTGTGCCCTAACATTC 38099
QY 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 38100 TGGATCGGTGATATCATTTGTGACATACACCCCTCCCGTATGCTCAACGCCCATTTGACTC 38159
QY 753 AlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeu 772
Db 38160 GCTGAGTGTTTGGAATCTACGCAAGTGTGCTGCTTCTTCTGTTGTCTCATG 38219
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
Db 38220 AAGTGCCGAGACAAAGGGCATGCTCTTGAAGTCATCACCGAGTCTTCTGTGCGA 38279
QY 793 AlaLysGlnAla 796
Db 38280 GCAAAAGCAGGCC 38291
RESULT 10
AE017116 300957 bp DNA linear PUN 06-JUN-2003
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 10, section 70 of 77 of the complete sequence.
ACCESSION
AE017116 AE016959
VERSION
AE017116.1 GI:31433303
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euphorbiaceae; Oryzae; Oryza.
REFERENCE
1 (bases 1 to 300957)
2 (bases 1 to 300957)
AUTHORS
Buell, C.R., Wing, R.A., McCombie, W.R., Messing, J. and Yuan, Q.
CONSTRM
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (05-MAY-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES.
source
1. .300957
location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
1. .12488
/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSUNBa0040D23 (GB:AC074196)."
3310. .>4569
/locus_tag="OSUNBa0040D23.23"
/note="similar to chloroplast nucleoid DNA binding protein CND41 GB:BA22813 GI:2541876 (Nicotiana tabacum)"
3310. .>4569
/locus_tag="OSUNBa0040D23.23"
misc_feature
gene
mRNA

CDS 3310. .4569
/locus_tag="OSJNBa0040D23.23"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54833.1"
/db_xref="GI:31433304"
/translation="MGRLEPMQLLVCLISVTTCAAAGLRRLDQGMGRRIADAT
AAPGGAVVPLHMSGAHYANFTIGTPQAVSGIVDSGELVMTQCAACRSSGCFKOE
LPEVDPASNTYRAEQCGSPICKSIPTRNCSGDGEGYEAPSMFGDTFGIASTDAI
GNAEGRLAFCGVASDGSIDGAMDGSGFVGLGRTPSLVGQSNVTAFSYCLAHGPG
KKSALFLGASAKLAGAKSNPPTLPDAYQALEKVTAAIGSPMANPEPFDLCTQNA
ASSGGAITVLQLETFRPLSYLPDAYQALEKVTAAIGSPMANPEPFDLCTQNA
VSGVPLVFTFQGGATLTADPSKYLGDGNGNGTVCLISLSTRLDSADDSILGSL
LQENVHFLFDLEKETLSFEPADCSLT"
<5409. .>5992
/locus_tag="OSJNBa0040D23.24"
/note="predicted by fgenesh"
join(<5409. .5547,5811. .>5992)
/locus_tag="OSJNBa0040D23.24"
join(5409. .5547,5811. .5992)
/locus_tag="OSJNBa0040D23.24"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54834.1"
/db_xref="GI:31433305"
/translation="MTGNLNSNIRICLTKTPTCLKHPCTGAMLRORTSTKAASHCP
TRGKRASKDSDLI SAVLDNQNP LNHGQTAKPLSTIVARGSTFVARRNPFKTRKKN
MYAL"
<8062. .>9183
/locus_tag="OSJNBa0040D23.25"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<8062. .>9183
/locus_tag="OSJNBa0040D23.25"
8062. .9183
/locus_tag="OSJNBa0040D23.25"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54835.1"
/db_xref="GI:31433306"
/translation="MRGLLDAAVAGGGGVPIYWSQPLYMANLTIGTPPOASAI
IHLGEFVWTQCSPCRCKFQDPLFNRSASSTRPEPCGTALCESVPASTCSGDVC
SYEETMFGDTSIGGTFPTAIGTATASLAFGAMDSNIKOLIGASGVGLGRTPSL
VGOMNATAFSYCLAPHGAAGKKSALLLGASAKLAGKSAATPLVNTSDSSDYMIHL
EGIKFGDVI IAPPPNGSVLVDTI FGVSTLVDAFOAIKAVTVAVGAAPMATPTKPF
DLCPKAAAAGANSLLPLDVLVTFQGAALTYPPSKYMYDAGNGTVCLAMSSAML
NLTTLSILGRHQENIHFLFDLDEKETLSFEPADCSLS"
<11356. .>12573
/locus_tag="OSJNBa0025B05.4"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<11356. .>12573
/locus_tag="OSJNBa0025B05.4"
11356. .12573
/locus_tag="OSJNBa0025B05.4"
/codon_start=1
/product="putative chloroplast nucleoid DNA binding
protein"
/protein_id="AAP54836.1"
/db_xref="GI:31433307"
/translation="MGRPVATLFLCFISVTACSLSEQATRGRLLAGVATPPAAGA
VAVPIYLSQGLYANFTIGTPQVSAVVDLTGELVMTQCTPCQCFEEDLPLFDP
KSTFRGLPCGSHLCESIPSSRNCTSDVCIVEAPTKAGDTGKAGTDFAI GAKE
LGFGCVMTDKRLKLTIGPSGIVGLGRTPSLVQMNVTAFSYCLAGKSGALFLGAT
AKOLAGKNSSTPEVITKTSAGSSDNGSNPYVMVKLAGIKTGAPLQAASSGSTVLLD
TVSRASYLADGAYKALKALTAAGVQPVASPPKYDLCPKAVAGDAPELVFTFDGG
AALTVPANVYLAAGNCTVCLTIGSSASINLTGBLGASILGSLQENVHFLFDLKEE
TLSFKPADCSSLT"
12489. .23141
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBa0025B05
(GB:AC096782)."
complement(<16252. .>17307)

mRNA /locus_tag="OSJNBa0025B05.3"
/note="predicted by fgenesh"
complement(<16252. .>17307)
/locus_tag="OSJNBa0025B05.3"
complement(16252. .17307)
/locus_tag="OSJNBa0025B05.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54837.1"
/db_xref="GI:31433308"
/translation="MVSAQAQGVAAACVTLFVPMGLAGMHLGRNKVLFSSGALFVSL
AVGVHLSPLYPLSLPHLLLASFHPHPISSSSSSSAASSCVPLHRVSWADAGESV
GRAMWPPSLASTCGLARLSRDDASILLNGSWVWAGDSQARLLVALLRLLLDPAA
AAAEPELFRHSDYRATVPARGISVDFVWAPFESNLTRLHEDLRAPRTPDVLVLS
GLMMLHVTDAARYGDALASVDAKSLRSPLPVPPHMFVLGLPLVHMLNTDAK
VHMNDTILQAYDLEVEQRGLQRDGPFLLLDVGKLSRGCGQCTADGMHYDGDVYDA
VLHIMNALVIESQRI"
complement(<18878. .>19249)
/locus_tag="OSJNB0064P21.1"
/note="similar to histone H4 (H4C13) GI:168499 (Zea mays),
(Plant Physiol Biochem 25, 235-247 (1987)); EST D23040
from this gene"
complement(<18878. .>19249)
/locus_tag="OSJNB0064P21.1"
complement(18878. .19189)
/locus_tag="OSJNB0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAP54838.1"
/db_xref="GI:31433309"
/translation="MSGRGKGKGLGKGAKRHKVLRDNIQTGTPAIRRLARRGV
KRISGLIYEETRGVLKIFLENVIRDAVYTEHARKKVTAMDVYALKRQRTLYFGG
G"
<19395. .>20598
/locus_tag="OSJNBa0025B05.2"
/note="similar to B regulatory subunit of PP2A
GB:AAB58900 GI:2160690 (Arabidopsis thaliana)"
join(<19395. .19473,19781. .19910,19992. .>20598)
/locus_tag="OSJNBa0025B05.2"
join(19395. .19473,19781. .19910,19992. .20598)
/locus_tag="OSJNBa0025B05.2"
/codon_start=1
/product="putative protein phosphatase 2A subunit"
/protein_id="AAP54839.1"
/db_xref="GI:31433310"
/translation="MTWLVTYGRDAIRGRVVLHRMWRMDGAPPGDGAGQDGRROL
VPASGQCRRPPTRRPRRPTCSTRTRLRNHVDRSFLAGLVALFAFEDSRERDLKY
YHQLSKLIVERAFMQRIMAAALRFVYDTSPEEAERHCGVGLLEICGSIINGFAVP
LKEHRAFMARVLLPLHRTRWVHYTRQLSYCGSIRSRHLLPPLAGDQMPEGAAHR
RARGDRDPLSAPGLPAASAAAHRYVCFHFHNRHFHSQKIDFIYICTFN"
23142. .155358
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNB0064P21
(GB:AC073166)."
<27557. .>31580
/locus_tag="OSJNB0064P21.2"
/note="similar to calcium dependent protein kinase
GI:435466 (Oryza sativa) (Gene 129 (2), 183-189 (1993));
EST C22538, C22540, AA750452, C99271, C22539 from this
gene"
join(27557. .28248,28582. .28725,28808. .28960,29138. .29253,
29890. .30054,30356. .30577,31147. .31580)
/locus_tag="OSJNB0064P21.2"
join(27567. .28248,28582. .28725,28808. .28960,29138. .29253,
29890. .30054,30356. .30577,31147. .31269)
/locus_tag="OSJNB0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAP54840.1"
/db_xref="GI:31433311"
/translation="MGNSQNGTPVGNVQNSNRFQNDRFASRYVDGNTEDCYSGSSR
ASLAGALROGLNLSKSPVLGYKTPVAVRELYTLGRELGGQGFQKTYLCTEISTGCOYACK
TILKSNLRCSVSDIEDVREIQLIMHLSGQKNIVITIKDYEDBQAVHIVMELCAGGELF

SKIOKRGHYSEKXAELIKIIVGIIETCHSHGVHRDLKPENFLLDADDFSVKAD

Alignment Scores:

Pred. No.:	1.78e-155	length:	300957
Score:	2514.00	Matches:	552
Percent Similarity:	53.77%	Conservative:	75
Best Local Similarity:	47.34%	Mismatches:	109
Query Match:	61.69%	Indels:	431
DB:	8	Gaps:	9

US-10-051-909-32 (1-800) x AE017116 (1-300957)

```
QY      56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db      48060 ATGGCGGGCGCCGTCGTGTCGCCATCGCGCCCTCCATCGGCAACTTGCTGCAGGGCTGG 48119

QY      76 AspAsnAlaThrIleAla-----
Db      48120 GATTAATGCAACCATTTGACAGTAAGTTCGTGTTAGAGAAAAAAGCAATCTTTTCTTCT 48179

QY      81 -----
Db      48180 GCCATAGTTTGTGTCATGTCGTATGTAGACGCTAGGCGCCGTGTTGTGCGCAATTAT 48239

QY      81 -----
Db      48240 TGTGATACTTGTGAGATCATGATTTGTTGTAATCATCTGCAGTAAATAAGTTT 48299

QY      81 -----
Db      48300 AGTTGTTCTTCTGTCATGATATCACAAAAAAACAACGTATATTTTTCAGTA 48359

QY      81 -----
Db      48360 GAATTTTGAGTCAATTATGATGAGATTTAACAATGTACGATCTCATGATGATCTGT 48419

QY      81 -----
Db      48420 ATTTATGGAACCTGATACTGTCTGCATCGTAAGGTTTCAGAAATTGATTGATACCGATG 48479

QY      82 -----
Db      48480 GTTTGACTTCTTGCTGACCCCTGTGACTTGTGGGATGATCTTTATCAGGTGCGTACT 48539

QY      85 uTyrIleIysIysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSe 105
Db      48540 GTACATCAAGAAGGAATTCAACTTGCAGAGCGAGCCCTTATTCGAAGCCCTGATCGTGC 48599

QY      105 rMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerI 125
Db      48600 CATGTCGTCATTGGGGCGAGCATCATCAGACGTTCTCTGAGCAGATGGCTGATCTTT 48659

QY      125 eGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe 145
Db      48660 TCGTAGGCGGCCCATGCTGATCGCGCTGCTCTCTACTTTGTTAGTGGCTAGTGAT 48719

QY      145 tLeuTyrSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyI 165
Db      48720 GCTTTGGGCGCCAAATGTATGTGTGCTCTTGCGAGGCTCATTGACGGGTTCCGGAT 48779

QY      165 eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr 185
Db      48780 CGGTTTGGCTGTCACGCTGTATCCATTGTACATCTCTGAGACTGCCCGCAGACATCAG 48839

QY      185 gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCy 205
Db      48840 AGGACTGCTTAACAACGCTGCCGAGTTCACTGGGTCTGAGGGGATGTTCTTTCATACTG 48899

QY      205 sMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMetLeuGlyValIe 225
Db      48900 CATGGTATTTGGCATGTCCCTCATGCCACAGCCAGATGGAGATCATGCTTGCGCTTCT 48959

QY      225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProAr 245
```

```
Db      48960 ATCATATCCATCACTTATATCTTTGCATTGACCATCTTTACTTACCTGATCGCCGAG 49019

QY      245 gTTPLeuValSerIyegIyArgMetAlaGluAlaIySlyValLeuGlnIySlyLeuArgG 265
Db      49020 GTGGCTCGTGAGCAAGAGAATGGCTGAGGCCAAGCGTGTGTGCAAGGCCCTCGCTGG 49079

QY      265 yLysAspAspValSer-----
Db      49080 AAGAGAAGATGTTTCAGGTTGTATGTCTTACCAGCATTTTCCCTGTGAGATCTAGA 49139

QY      270 -----
Db      49140 AAGTGAGAACTATTATGCTATACGATAGATATATATCTATGCGATGAGTTTCT 49199

QY      271 -----
Db      49200 CATGTTGAATTACTGTGCAGAGAAATGGCCCTTCTGTTGAAGGCTGCGGGCTTGCGA 49259

QY      284 lYAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuV 304
Db      49260 AAGACACAAAATTGAGGAATACATAATTGACCTGATGATGAGCTTGCTGATGAAAGGCC 49319

QY      304 aThrAspGlyAspIySlyGluIleThrLeuTyrGlyProGluGluGlyGlnSerTyrPi 324
Db      49320 TGGCTCCA--GATCAGAGAAAGATCAAACTGTATGCTCTGAAGAGGCTTATCGTGGG 49376

QY      324 lAlaArgProSerIySlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH 344
Db      49377 TTGCCCGTCTGTTCACGGGCAAAAGTGACCTTGAAGTGCAATTAGTCTCATCTCTCGTC 49436

QY      344 lSgIySerMetValAngIlnSerValProLeuMetAspProIleValThrLeuPheGlyS 364
Db      49437 ATGCTAGTATGTGTCAAGGTAAGGCCCTGTGTGATCTCTGTGACCCCTTTTGGA 49496

QY      364 eValHisGluAsnMetProGlnAlaGlyIySerMetArgSerThrLeuPheProAsnP 384
Db      49497 GTGTCCATGAGAGATGCTGAGATGATGGAAGCATGCGGAGACACATTTTCTTA 49556

QY      384 heGlySerMetPheSerValThrAspGlnHisAlaIySAsnGluGlnTyrAspGluGlu 404
Db      49557 TTGGCAGCATGTTTGTGTGGCGGAACAGCAGCAAGCTAAAGTGATTTGGATGCTGAGA 49616

QY      404 snLeuHisArgAspArgGluIyTyrAlaSerAspGlyAlaGlyIyAspTyrGluAspA 424
Db      49617 GT---CAACGGAGGGGTGAAGATTATGATCAGACCATGTTGGGATGACATGAAGATA 49673

QY      424 snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyIyAspIle---V 443
Db      49674 GCCTCCAAGGCCACTTATTCTCTGTCAGCGCAACAGCGTGAAGGAAAGAGATCGCTG 49733

QY      443 alHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu 463
Db      49734 CACCTCATGGCAGTATATGCGTGTGGA-----AGAACTAGTACTTCATGTCAGG 49787

QY      463 lYgIyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTyrIySTrs 483
Db      49788 GCGGGAGGAGAGTAAGCAGATGGGCAATGGGAGATGGCAGATTGGCTTGAATAATGA 49847

QY      483 eGlyIySgIyGluIyAsnGlyArgIySgIyGlyIyPheIySArgValTyrLeuHisG 503
Db      49848 CTGAGAGAGAAAGGTGAGATGGCGAAAAAGAGGTGGCTTCCAACGTATCTATCTGCATG 49907

QY      503 lngIyGlyValProGlySerArgArgGlySerIleValSerLeuProGlyIyGlyAspV 523
Db      49908 AAGAGGGGTGTGACAGGTGATCGCAGGGGCTCTATCTCATTTGCCCT--GGAGGTGATG 49964

QY      523 alPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheS 543
Db      49965 TTCCTCCTGTGTGTGAGTTCCTCCAGGCAAGCTCTTGTACGCCAACCTGCTTTTACT 50024

QY      543 eRlySgIyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVal 563
```

Db 50025 CTAAGGAATTGATGAGCAACGCCCTTGCTGGCCCTGCTAAGTGCAATCCATCTCAGGAG 50084
QY 563 laAlaIysGlySerArgTrpIysAspLeuPheGluProGlyValArgAlaLeuLeu 583
Db 50085 TTGCTAAAGTCCAAAATGGGAGCACTTATTGCAACCTGGAGTGAACATGCTCTGTTG 50144
QY 583 alGlyValGlyIleGlnIleuGln----- 591
Db 50145 TTGGCATAGGATACAAATCCCTGCAACAGGTAACCCTTTTCGATTTTATGTCATGTAC 50204
QY 591 ----- 591
Db 50205 GAATGCATGTCTATGTGTGAACCTAATTTTCTAGAGCTCCAGATAAATTACTGTGAT 50264
QY 591 ----- 591
Db 50265 TAAATGAAGCTGTTAAATCAAAGCATCTTCATGCAAGTGTCTTCTAGGTTCTAGCGAG 50324
QY 591 ----- 591
Db 50325 TCCTCATTTAAAGCTCTTACTTACTTACTCGAAAACAGTTTAGCTGTACGTTT 50384
QY 591 ----- 591
Db 50385 TCCTGCATATATATTCATACAAAGCAATGAGCTAAACTGCACAATAATACACCTTG 50444
QY 591 ----- 591
Db 50445 ATGATTAGAATACAATGCCATCAATGTTGTTAGTTTTCACCTAGTAAGTCAACT 50504
QY 591 ----- 591
Db 50505 TAAATCTTGCAGTATTCATTATTAGAAGAAACCATTGGAGCTGGGCTATGTCACATGACC 50564
QY 591 ----- 591
Db 50565 TGAATTACGGGTATTCTATCTGAAGCAGAGAAACAAATGTCATGCTTTTATTCCTG 50624
QY 591 ----- 591
Db 50625 AATTTATTGCTGAATATGTTCTAGTTTCGAGAAATAACTTTCAGAACTTTTCACAGT 50684
QY 591 ----- 591
Db 50685 GGCAATATCATGTATATCTCTTTTACTTCCCTCCACTCTTCTATTTTCCCTGTAT 50744
QY 591 ----- 591
Db 50745 AATTATATCAGTCTATTCATGTTTACCTCAGTATAGCATCTGAATGTTAATGTGCTAT 50804
QY 592 -----GlnPheAlaGlyIleAsrGlyValLeuTyrTyrThrProGln 605
Db 50805 CAACTTTTCCATTATGTCAGTTTGCTGGCAATTAATGAGTTCTGTACTACACTCCCAA 50864
QY 606 IleuGluGlnAlaGlyValAlaValIleuSerLysPheGlyLeuSerSerAlaSer 625
Db 50865 AATCTTGAGCAAGCTGTTGTTGTTCTTCTTGCAACATTTGACTTAGCTCCTCATCT 50924
QY 626 AlaSerIleuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
Db 50925 GCATCTATTTCTTATTAGCGAGTGAACAACCTTGCTGATGCTTCCAGCATGTGTAATGCT 50984
QY 646 MetLeuLeuMetAspLeuSerGlyArg----- 655
Db 50985 ATGAGGCTCATGATATGTCTGAGAGAAGGTAGACTTCATATTTTACATTTTCTTGT 51044
QY 655 ----- 655
Db 51045 GATTACATGCATCACTTCTAAGTGAAGATATTATGCTAACCATTAATTCGCGTTGTTCT 51104
QY 656 -----PheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 51105 CATTCAGGTTTCTT-CTCCTTGCAACAATCCCTATCTGATAGTACACTAGCTATCTTG 51163

QY 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
Db 51164 ATTCTGTCATATATTCTGATGTGGGACCATGGTTTCATGCTCCTCACTGTCCAGTCACT 51223
QY 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 51224 GTCATACTCTACTTCTGCTTCTTGTGATGGGTTGGGCTTATTCAAACATTCCTGT 51283
QY 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db 51284 GCAGAGATTTCCCGACCAACCGTTCGTGGCAGTGCATAGCCATCTGTGCCCTAACATTTC 51343
QY 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 51344 TGGATCGGTGATATCATTTGTGACATACACCCCTCCCGTATGCTCAACGCCATTGACTC 51403
QY 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db 51404 GCTGGAGTGTGGAAATCTACGCAAGTGGTGTGCATACCTGCTTCTGTTGTTCTCATG 51463
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
Db 51464 AAGTGCCGGAGACAAAGGCGATGCTCTTGAAATCATCACCAGAGTCTTCTGTGCGGA 51523
QY 793 AlaLysGlnAla 796
Db 51524 GCAAAAGCAGGCC 51535
RESULT 11
ATF23E12 86710 bp DNA linear PLN 01-APR-1999
LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone F23E12 (ESSA
DEFINITION project).
ACCESSION AL022604 GI:3080406
VERSION AL022604
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Bevan, M., Hilbert, H., Braun, M., Holzer, E., Brandt, A.,
Duesterhoeft, A., Hoheisel, J., Mewes, H.W., Mayer, K.F.X. and
Schueller, C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 86710)
AUTHORS EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (01-APR-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
schueller@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES
source
1. 86710
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
173. .1890
/gene="F23E12.10"
join<173. .374,456. .919,1089. .1610,1678. .1890)
/gene="F23E12.10"
/codon_start=1
/product="hypothetical protein (fragment)"
/protein_id="CAA18726.1"

/db_xref="GI:3080407"
/db_xref="SPTREMBL:O65486"
/translation="FNAAVAGCFPAVESVIMPSSSTDSTSLPNTTSMVILISAVTASY
VSEIGLGSPEAFKVPDYDFERSPEGLPLYLLGALCVSLALSRCTSSMTSAVDSLK
DAGIPKAVFPVMGSLVGIILALVPEVLTYGFQNVLDILLEKRPFKGISADLLQLVA
VKIAATAMCRASGLVGYYAPSLFIGAGMAYGKEIGLALQNPDPNLSILEVASPQ
AYGLVGMATLAVGCQVPLTAVLLFELTDYRIVLPPLGAVGMSSWTSGSKROET
RETKETRKRKSGEAVQSLTSSDDESSNNLCVESSSLCLDDSLNQSLELPKSI FVSEA
MTRFATVMSTSLLEALTRMLIEKQSCALIVDPDNIPLGILTLSDIOEFSKARKEGN
NRPKDIFVNDICSRSGKCKVPWTVPMDLLAAQTIMNKHELSHAVVSGSIDAPRI
HPVGLDRECTILTRR"
exon 173. .374
/gene="F23E12.10"
/number=1
intron 375. .455
/gene="F23E12.10"
/number=1
exon 456. .919
/gene="F23E12.10"
/number=2
intron 920. .1088
/gene="F23E12.10"
/number=2
exon 1089. .1610
/gene="F23E12.10"
/number=3
intron 1611. .1677
/gene="F23E12.10"
/number=3
exon 1678. .1890
/gene="F23E12.10"
/number=4
gene 3124. .4428
/gene="F23E12.20"
join(3124. .3241,3323. .3465,3595. .3696,3945. .4107,
4211. .4428)
/gene="F23E12.20"
/note="similarity to dihydroflavonol-4-reductase, Gerbera
x sp., PIR2:S35189"
/codon_start=1
/product="putative protein"
/protein_id="CA18727.1"
/db_xref="GI:3080408"
/db_xref="SPTREMBL:O65487"
/translation="MDQAKGKVCVTGASGFLASWLVRLLLEGYEVIGVVDPGNEKK
LAHLWLKLEGAKERLRVKADLMEEGSPDNAMGCGVFHTASPEILRAIEGTINLIR
SCRKNPSLKRIVLTSSSTVWYALSKTLAEQAAMKFSBENGIDLVYLPSPFLVGPSP
PDLCTASDVLGLKGETEKFQWHGOMGYHIDVARTHIVVFHEAAQGRYICSSNV
ISLEELVSFLSARYPSLPPIPKRFTDPLN"
exon 3124. .3241
/gene="F23E12.20"
/number=1
intron 3242. .3322
/gene="F23E12.20"
/number=1
exon 3323. .3465
/gene="F23E12.20"
/number=2
intron 3466. .3594
/gene="F23E12.20"
/number=2
exon 3595. .3696
/gene="F23E12.20"
/number=3
intron 3697. .3944
/gene="F23E12.20"
/number=3
exon 3945. .4107
/gene="F23E12.20"
/number=4
intron 4108. .4210
/gene="F23E12.20"
/number=4

exon 4211. .4428
/gene="F23E12.20"
/number=5
gene 4939. .6163
/gene="F23E12.30"
complement(join(4939. .5013,5109. .5195,5307. .5342,
5451. .5651,5725. .5811,6161. .6163))
/gene="F23E12.30"
complement(join(4939. .5013,5109. .5195,5307. .5342,
5451. .5651,5725. .5811,6161. .6163))
/gene="F23E12.30"
/note="Contains Clathrin adaptor complexes small chain
signature [VYKRYASLYF]
contains EST gb:T44219"
/codon_start=1
/product="clathrin assembly protein AP19 homolog"
/protein_id="CA18728.1"
/db_xref="GI:3080409"
/db_xref="SPTREMBL:O23685"
/translation="MIHFVLVSRQGVRLTKWYSPYAQKERSKVIRELGVILNRGP
KLCNFEVWRGYKVVKRYASLYFCNIDQEDNELEVEIHHYVEILDRYFGSVCELD
LIENFHAYYIILDELLIAGELQESSKKTVARIIISAQDQLEVAKEAASSISNIIAQAT
NR"
complement(4939. .5013)
/gene="F23E12.30"
/number=1
intron complement(5014. .5108)
/gene="F23E12.30"
/number=1
exon complement(5109. .5195)
/gene="F23E12.30"
/number=2
intron complement(5196. .5306)
/gene="F23E12.30"
/number=2
exon complement(5307. .5342)
/gene="F23E12.30"
/number=3
intron complement(5343. .5450)
/gene="F23E12.30"
/number=3
exon complement(5451. .5651)
/gene="F23E12.30"
/number=4
intron complement(5652. .5724)
/gene="F23E12.30"
/number=4
exon complement(5725. .5811)
/gene="F23E12.30"
/number=5
intron complement(5812. .6160)
/gene="F23E12.30"
/number=5
exon complement(6161. .6163)
/gene="F23E12.30"
/number=6
gene 7352. .7582
/gene="F23E12.40"
complement(7352. .7582)
/gene="F23E12.40"
complement(7352. .7582)
/gene="F23E12.40"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA18729.1"
/db_xref="GI:3080410"
/db_xref="SPTREMBL:O65488"
/translation="MPLVRKNGVWRIGSPMLHSESNLVGDGVAPPRGERDEIGEYV
GGETINVSSESCGEEGAQDAKTPEGEADRI"
complement(7352. .7582)
/gene="F23E12.40"
/number=1
gene 8233. .9045

gene /gene="F23E12.50"
complement (8233 . 9045)
/gene="F23E12.50"
CDS complement (8233 . 9045)
/gene="F23E12.50"
/note="similarity to DNA-binding protein PDI, Pisum sativum"

/codon_start=1
/product="putative protein"
/protein_id="CA18730.1"
/db_xref="GI:3080411"
/db_xref="SPTREMBL:O65489"

Alignment Scores:
Pred. No.: 1.86e-155 Length: 86710
Score: 2502.00 Matches: 523
Percent Similarity: 68.91% Conservative: 91
Best Local Similarity: 58.70% Mismatches: 124
Query Match: 61.40% Indels: 153
DB: 8 Gaps: 12

US-10-051-909-32 (1-800) x ATF23E12 (1-86710)

QY 52 LeuGIuAsPlys--MetSerGIyAlaValLeuValAlaIleValAlaSerIleGIyAsn 70
Db 39649 ATAGAAGACGGAATTATGAGTGAAGCTGTGCTTGTCTAATGCTGCTGTGGCAAC 39708
QY 71 LeuLeuGIingIYTrpAspAsnAlaThrIle----- 80
Db 39709 TTGTTACAAGATGGGATTAACGCACTATTGCAGTGATGATTATCTTCTGTATCCATA 39768
QY 80 ----- 80
Db 39769 ATTCTGTTTATGAATTAGATCGATAAAGTCAGATTTTATGATACATTCACAGA 39828
QY 81 -----A1 81
Db 39829 TCTTTTAGCTTATCTCTTTGCTGTGTGATGATTTTAACTTTCTTGGGTACTCTGC 39888
QY 81 aAlaAlaValLeuTyrIleYsYsGIuPheGIuLeuGIuAsnGIuProThrValGIuGI 101
Db 39889 AGGAGCTGTGTGTACATAAAGAGATTAAATTGGAGTAATCCATCAGTGAAGG 39948
QY 101 yLeuIleValSerMetSerLeuIleGIyAlaThrIleValThrThrPheSerGIyProle 121
Db 39949 TCTAATTGTGGCGATGTCACTTATGTGTCTACTGTATTACAACATGCTCTGGAGGGT 40008
QY 121 uSerAspSerIleGIyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSe 141
Db 40009 AGCTGATTTGGCTTGTGCGCCGCCATGCTAATATGTCTCAATCTCTACTTTGTGG 40068
QY 141 rGIyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuLeuAlaArgPheValAs 161
Db 40069 TTCTCTAGTAATGCTATGCTCCGAATGTTATGTGTGCTCTTAGGAAGGTTGTAGA 40128
QY 161 pGIyPheGIyIleGIyLeuAlaValThrLeuValProLeuTyrIleSerGIuIleAlaPr 181
Db 40129 TGGATTGGGGTGTGCTTGTGTGCTACACTTGTTCCTATTATATATCTGAGACTGCACC 40188
QY 181 oSerGIuIleArgGIyLeuLeuAsnThrLeuProGIuPheSerGIySerGIyGIyMetPhe 201
Db 40189 ACCTGAGATTAGGGGACTGTGAATAACGCTACCGCAGTTCACTGGCTCTGAGGGATGTT 40248
QY 201 eleuSerTyrCysMetValPheGIyMetSerLeuSerProSerProAspTrpArgIleMe 221
Db 40249 CTATCTTACTGATGTTTCGGAATGTCGTTGATGCCATCACCTAGCTGAGATTGAT 40308
QY 221 tLeuGIyValLeuAlaIleProSerLeuPhePhePheGIyLeuThrIlePheTyrLeuPr 241
Db 40309 GCTTGTGTCTTTTCATCCCTTCCCTTGTCTTTCTTCCTACAGGTCTTCTTGTGCC 40368
QY 241 oGIuSerProArgTyrLeuValSerIySGIyArgMetAlaGIuAlaYsYsValLeuGI 261

Db 40369 CGAGTCCCCAAGGTGGCTCGTGACCAAGGTGCAATGCTTGAAGCAAAAGCGGTCTTCA 40428
QY 261 nLyLeuArgGIyLyAspAspValSer----- 270
Db 40429 GAGACTGCGTGTGCGGAAGATGTGTCTGTAGTCTTAATCGTTCTGACTTAATTCT 40488
QY 271 -----GlyGIuLeuS 274
Db 40489 TTCATGAACCAATCTTTCTATTGAATGTTCTTATGTGCAATGTATAGGTGAGATGG 40548
QY 274 erLeuLeuLeuGIuGIyLeuGIuValGIyGIyAspThrSerIleGIuIuTyrIleIleG 294
Db 40549 CTTTGTGTGGGCTCTTGAATTGGAGTGAAACACCATAGAGAAATATATAATTG 40608
QY 294 lYProAlaThrGIuAlaAlaAspAspLeuValThrAspGIyAspLyGIuGIuIleThrI 314
Db 40609 GTCCCGCGATGAAGTTACTGATGATCATGATATAGCTGTGATTAAGATCAAAATTAAGT 40668
QY 314 eUTyrGIyProGIuGIuGIyGIuInSerTrpIleAlaArgProSerIySGIyProIleMetL 334
Db 40669 TATATGTGCAGAGAAGGGCTGAGTTGGTGTCTAGCCAGTCAAGGA----- 40718
QY 334 euGIySerValLeuSerLeuAlaSerArgHISGIySer---MetValAsnGIuSerValP 353
Db 40719 --GGAAGCACTATGAGTCTTTGTCTCGCCATGGAAGTACATGACGACGCAAGCT 40776
QY 353 roLeuMetAspProIleValThrLeuPheGIySerValHISGIuAsnMetProGIuAlaG 373
Db 40777 CATGTATGATCTCTTGTCTCACACTGTTTGGAGCGCTTACGAGAAGATGCCGACACT- 40835
QY 373 lYGIySerMetArgSerThrLeuPheProAsnPheGIySerMetPheSerValThrAspG 393
Db 40836 --GGAAGCATGAGAGTCCCTTGTCCACATTTTGGAGTATGTTGAGTGTGGAGGGA 40893
QY 393 lnhISAlaYsAsnGIuGIuInTrpAspGIuAsnLeuHISArgAspAspGIuGIuTyrA 413
Db 40894 ATCAACCAAGACATGAAGATTGGATGAAGAACTTGTGGAGAAGGTGAGGATTATC 40953
QY 413 lAsErAspGIyAlaGIyGIyAspTyrGIuAspAsnLeuHISerProLeuLeuSerArgG 433
Db 40954 CATCCGAC--CATGGAGATGATCTGAAGATGATCTTCATTTCCGTGATCTCACGTC 41010
QY 433 lnaIaThrGIyAlaGIuGIyLyAspIleValHISGIyHISArgGIySerAlaLeuS 453
Db 41011 AAACGACACGATGAG--AAAGACATGCCCTCACACTGCTCATGAACTTTTCTACCT 41067
QY 453 erMetArgArgGIuInThrLeuLeuGIyGIuGIyAspGIyValSerSerThrAspIleG 473
Db 41068 TCAGACATGGAAGTCAAGTCAAGGAGCTCAAGGGGAAGGAGCGGTACTATGGGATTG 41127
QY 473 lYGIyGIyTrpGIuLeuAlaTrpLySTrpSerGIuLySGIuGIyGIuAsnGIyArgLySG 493
Db 41128 GAGGTGATGGCAAGTGGCATGGAATGGAAGAAAGAGAAGATGAATCGGACAGAAAG 41187
QY 493 lGIyGIyPheLyArgValTyrLeuHISGIuGIuGIyValProGIySerArgArgGIyS 513
Db 41188 AAGGTGGTTTAAACGATATACTTGCAACAAGAGTTTCCAGGATCTCGACGTGGCT 41247
QY 513 erIleValSerLeuProGIyGIyGIyAspValPheGIuGIySerGIuPheValHISAlaA 533
Db 41248 CAATTGTTCAATGCCCTGCTGTGATGAACCGGTGAG--GCAGATTTGTACAAAGCGT 41304
QY 533 laAlaLeuValSerGIuInSerAlaLeuPheSerIySGIyLeuAlaGIuProArgMetSerA 553
Db 41305 CTGCTTTGTTTAAACCAACGAGCTCTTATTTCCAAAGACCTTCTCAAGACATACATTTG 41364
QY 553 spAlaAlaMetValHISProSerGIuValAlaAlaYsGIySerArgTrpLyAspLeuP 573
Db 41365 GTCCTGTATGTATCATTCATCCGAA--ACAACATAAAGGGTCAATTGGCATGATCTTC 41421
QY 573 heGIuProGIyValArgArgAlaLeuLeuValGIyValGIyIleGIuIleLeuGIu---- 591
Db 41422 ATGATCCTGAGTCAACGCGTGCATTAAGTGAAGAGTTGACCTTCAATACTTCAGCAGG 41481

QY	591	-----	-----	591
Db	41482	TAGTCTAGATAGCCACATCAAAATTAGTAATCAGAGAATAATATGCGGTTATGCTA	41541	
QY	592	-----	-----	601
Db	41542	ATTCTATGTTTCCGCTTCCTCTTGGAATGCAAGGCTTCAGGCATCAACGAGTTCTTT	41601	
QY	601	YRTYRThrProGlnIleLeuGluGlnAlaValIleLeuSerLysPheGlyL	621	
Db	41602	ACTACACACCGCAATCTTGAGCAGCGGGTGTGGGATCCTACTATCGAATGGGA	41661	
QY	621	euserSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProc	641	
Db	41662	TTAGTTCTTCCTCAGCATCTTACTTATTAAGTCATTGACAACTTTGTGATGTTACCTG	41721	
QY	641	ysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg-----	654	
Db	41722	CAATAGCTGTTGCAATGAGGCTCATGATCTTCTGTGGAAGTGATTAACCTTAAT	41781	
QY	654	-----	-----	654
Db	41782	TTCCCTTGTCTTACATACATTAACGAAACAGCTAAATGCGTTATGAATGTTCTT	41841	
QY	655	-----	-----	669
Db	41842	TTATTTTATATTGGCAGACCTTGCTTCTCACCACGATACCAATCCGTAGATCTCTTA	41901	
QY	670	ValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuSer	689	
Db	41902	TTGTTTATGTAATCTCAATCTTGTTCACATGAACAGCATTTGTGACGCGGCTTATCA	41961	
QY	690	ThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsn	709	
Db	41962	ACCGTAAGCGTTGCTGCTACTTCTGCTTCTGATGGGTTTCGGTCCGCTCCAAAC	42021	
QY	710	IleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAla	729	
Db	42022	ATCCTCTGTTCAGAGATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCCA	42081	
QY	730	PheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAla	749	
Db	42082	CTCACCTTCTGGATCTGTGACATATATGCTCACTTACAGTCTCCCGTGTGCTCAAAATCC	42141	
QY	750	IleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPhe	769	
Db	42142	ATTGAGACTAGCTGTGTGTTGGAATGACGCAATCGTATGTGCATTTCAATGGGTTCTT	42201	
QY	770	ValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe	789	
Db	42202	GTGTTTCATTAAGTCCCGGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCTT	42261	
QY	790	AlaValGlyAlaLysGlnAlaAlaAla 798		
Db	42262	TCTGTTGAGCTAGACAGCTGAAGCT 42288		
RESULT 12				
ATCHRIV83/c				
LOCUS	ATCHRIV83	197859 bp	DNA	linear
DEFINITION	Arabidopsis thaliana	DNA chromosome 4, contig fragment No. 83.		
ACCESSION	AL161587			
VERSION	AL161587.2	GI:7270470		
KEYWORDS				
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 24256)			
AUTHORS	Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.			
JOURNAL	Unpublished			

REFERENCE	2 (bases 16641 to 103350)
AUTHORS	Hilbert,H., Braun,M., Holzer,E., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	3 (bases 98352 to 125759)
AUTHORS	Robben,J., Gymonprez,B., Volckaert,G., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	4 (bases 120761 to 197859)
AUTHORS	Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	5 (bases 1 to 197859)
AUTHORS	EU Arabidopsis sequencing project.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge laboratory, John Innes Centre, Colney lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV82 at the 5' end and an overlap with ATCHRIV84 at the 3' end.
FEATURES	Location/Qualifiers
source	1..197859
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
gene	11339..12705
	/gene="AT4g35190"
CDS	join(11339..11494,11536..11664,11853..11952,12119..12189,12314..12457,12499..12705)
	/gene="AT4g35190"
	/note="similarity to ctf, Mycobacterium leprae, gb:U15180-contains EST gb:T45691, T22640, T75954"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB80236.1"
	/db_xref="GI:7270471"
	/translation="MEIVKSRFRKVCVFCGSSGKRECYSDATDLAQLVRLCLNL ELENLKWVTRLNLVYGGSIGMLVSOAVHEAGHVGVAOIYDLFTLITGETYG EVIAYADMERKAEKMARHSDCFALPGYGTLEELLVIANADLGIDHPVGLNVDG YNNVLTFLDKAVDDGFIKPSORHIFVSAPNAKELVOKLEILINKDKKFDSDAIDLL FFDIPQAYKPVNDGVIAKSRWEVEKKVQQPQQQQVVFCSNTSMQTEIAL"
exon	11339..11494
	/gene="AT4g35190"
	/number=1
intron	11495..11535
	/gene="AT4g35190"
	/number=1
exon	11536..11664
	/gene="AT4g35190"
	/number=2
intron	11665..11852
	/gene="AT4g35190"
	/number=2
exon	11853..11952
	/gene="AT4g35190"
	/number=3
intron	11953..12118
	/gene="AT4g35190"
	/number=3
exon	12119..12189
	/gene="AT4g35190"
	/number=4
intron	12190..12313
	/gene="AT4g35190"
	/number=4


```

exon      12314. .12457  

          /gene="AT4g35190"  

          /number=5  

intron    12458. .12498  

          /gene="AT4g35190"  

          /number=5  

exon      12499. .12705  

          /gene="AT4g35190"  

          /number=6  

gene       13757. .14518  

          /gene="AT4g35200"  

          complement(13757. .14518)  

          /gene="AT4g35200"  

          complement(13757. .14518)  

          /note="similarity to various predicted proteins,  

          Arabidopsis thaliana"  

          /codon_start=1  

          /product="putative protein"  

          /protein_id="CAB80237.1"  

          /db_xref="GI:7270472"  

          /translation="MAVSFHVRSNSYPSRQHPQAHVDEQLTRLRSSGTAASSSSICQR  

          LSNLQDLHDSLEKMRILSVTNLALSDQIEKLDDSLRILDLCNIADAIISQKEGLM  

          EIQSILRRKPGDLSGEVKKYLVSRKFLKSLOKVIKSLKVCQSKDSTNALSIVFGRAE  

          AVTMAILFESLFSEFMSGKACGKWSLVSKMSQNKVTCAEANEFTRIDSEFQSEKSIQ  

          MEDVQNLIESCIQDLEDGISLSKSLIKYRVSILNI"  

          complement(13757. .14518)  

exon      /gene="AT4g35200"  

          /number=1  

gene       16043. .16795  

          /gene="AT4g35210"  

          16043. .16795  

          /gene="AT4g35210"  

          /note="similarity to various predicted proteins,  

          Arabidopsis thaliana"  

          /codon_start=1  

          /product="putative protein"  

          /protein_id="CAB80238.1"  

          /db_xref="GI:7270473"  

          /translation="MAVSFHVRSNSYPSRFQHPQAHVDEQLTRLRSSGTAAASSSICQR  

          LSNLQDLHDSLEKMRILSVTNQALSDQIEKLDDSIKILDLCSISKDGISQMKESLK  

          EIQSIVRRKRGDLSAENVKYYLASRKFLKSEFEKVLKSLKTSQNKNDAVFGAETVT  

          IALFESLFSEFMSGKACGKWSLVSKMSQSGKTCEAFANEFTRVDMEFQSEKSLQMED  

          VQNLERICQDLEDGISLSKSLIKYRVSILNI"  

exon      16043. .16795  

          /gene="AT4g35210"  

          /number=1  

misc_feature 16640. .24256  

            /note="position 76883-84499 overlaps to BAC clone F23E12,  

            EMBL acc:AL022604; for sequence analysis please refer to  

            this accession"  

            17252. .18625  

            /gene="AT4g35220"  

            join(17252. .17657,17740. .17784,18534. .18625)  

            /gene="AT4g35220"  

            /note="similarity to predicted protein, Arabidopsis  

            thaliana  

            contains EST gb:F14002, T46305, T43454, AI996658.1"  

            /codon_start=1  

            /product="putative protein"  

            /protein_id="CAB80239.1"  

            /db_xref="GI:7270474"  

            /translation="MAVPPLFLTLTLSPLLISAGASNAYPSIPGTAPIDGGFTDE  

            LKPIRREVYNGNKIYDISHRYPPEMPMDSEIGIFLMLAASMKNGSIANNSEMkip  

            THTGTHVDSRGHYVDKYDAGFDVDSLQLYLNGALLVDPKDKNTDWLEQGKLGLOF  

            VASSSIDSLPSQTQIRNCp"  

exon      17252. .17657  

          /gene="AT4g35220"  

          /number=1  

intron    17658. .17739  

          /gene="AT4g35220"  

          /number=1  

exon      17740. .17784

```

```

/gene="AT4g35220"
/number=2
17785. .18533
/gene="AT4g35220"
/number=2
18534. .18625
/gene="AT4g35220"
/number=3
19941. .22657
/gene="AT4g35230"
complement(join(19941. .20231,20345. .20458,20546. .20710,
20792. .20980,21340. .21425,21551. .21656,21755. .21888,
22078. .22213,22304. .22657))
/gene="AT4g35230"
complement(join(19941. .20231,20345. .20458,20546. .20710,
20792. .20980,21340. .21425,21551. .21656,21755. .21888,
22078. .22213,22304. .22657))
/gene="AT4g35230"
/note="similarity to protein kinase APK1, Arabidopsis
thaliana, PIR2:S28615
contains EST gb:A1997799.1, F13911"
/codon_start=1
/product="putative protein"
/protein_id="CAB80240.1"
/db_xref="GI:7270475"
/tranlation="MGCCQSLFSGDNPLGKDVQPQLPSQNHGCAITADNGSSGAS
GVGGGGGGGIPSFSEFPADLKAATNPFSSDNIIVSEGEKAPLVYGRLONRWIA
VKFKTKAMPEPKQFAEBAMVGKLHNRNLNLIGYCCDDGERLLVAEMPNDLAKH
LFHMENQTIEMAMRLVGYIAEALDYCSIEGRPLHDINAYRVLFDEGDPRLSCFG
LMKNSRDGSYSTNLAYTPPEYLNRGRVPESVTYSFGVLIDLGLSGKHIPPSHALDM
IRGNILILMDSHLLEGKSTEEATVVVELAQCLQEPRERPNTKDLVATLAPLOTKS
DVP SYVMLGIKKEEAPSTPORPLSPGACSRMDLTAIHQILVMTHYRDEGTNELS
FOEWTOQMCKMLDARKRGDQSFRKDFRTAIDCYSONSYGIIHCFCGPFDVGTWSP
TVFGRSLCYLLCDQPDALRDAMOQCYPDPWPTAFYMOSVALAKLNMTDAADMLN
EAAQLBEKRQRGGRGS"
BAALBEKRQRGGRGS"
complement(19941. .20231)
/gene="AT4g35230"
/number=1
complement(20232. .20344)
/gene="AT4g35230"

intron

Alignment Scores:
Pred. No.:      6.15e-155      length:      197859
Score:          2502.00        Matches:     523
Percent Similarity: 68.91%    Conservative: 91
Best Local Similarity: 58.70%   Mismatches:  124
Query Match:    61.40%       Indels:     153
DB:              8           Gaps:       12

US-10-051-909-32 (1-800) x ATCHRIv83 (1-197859)

QY      52 leugluasplys--MetSerGlYAlaValleuValAlaIleValAlaSerIleGlyAsn 70
Db      63702 ATAGAAGACGGAATTATGAGTGAGCTGTGCTTATGCTGCTGCTGTCGCAAC 63643
QY      71 leuleungllyTrpaaspasnaIatThrIle----- 80
Db      63642 TTGTTACAAGATGGGATAACCAACTATTGCAGGTGATGATTATCTTCTGATCCATA 63583
QY      80 ----- 80
Db      63582 ATTCTGTTATGAATTTAGGATCGATAAAGTCAGGATTTTATGATACAAATTCACAGA 63523
QY      81 -----Al 81
Db      63522 TCTTTAGCTTATCTCTTTTGGCTGTGTGANGCATTTTAAGTTTCTTTGGGTACTCTGC 63463
QY      81 aAlaAlaValleuTyrlIleLySLyglUpheGlnleuGlnasngluProThrValGluG1 101
Db      63462 AGGAGCTGTGTGTACATAAAAAAGAGATTAAATTGGAGAGTAATCCATCAGTGAAGG 63403
QY      101 yLeuIlleValSerMetSerleuIlleGlyAlaThrIleValThrThrPheserGlyProle 121

```

Db 63402 TCTAATTGGCGGATGTCATTAATTGGTGCTACTCTGATTAACAACATGCTCTGGAGGGCT 63343
QY 121 userAspSerileGlyArgArgPrometLeuileuSerSerileuTyrrPhePheSe 141
Db 63342 AGCTGATTGGCTTGCGCGCTCCCATGCTAATATGTCCTCAATTCTTACTTGTGG 63283
QY 141 rGlyLeuileMetLeuTrpSerProAsnValTyrrValLeuLeuAlaArgPheValAs 161
Db 63282 TTCTCTAGTAATGCTATGCTCTCCGAATGTTATGTTGCTCTTAGGAAGGTTTGA 63223
QY 161 pGlyPheGlyIleGlyLeuAlaValThrLeuValProleuTyrrIleSerGluileAlaPr 181
Db 63222 TGGATTGGGGTGGTCTTGTGCTCACACTTGTCTTATTAATATCTGAGACTGCACC 63163
QY 181 oSerGluileArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPh 201
Db 63162 ACCTGAGATTAGGGGACTGTTGAATAACGCTACCGCAGTTCACTGGCTTGAGGGATGTT 63103
QY 201 eleuSerTyrrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMe 221
Db 63102 CTATCTTACTGATGCTTTCCGAATGTCGTTGATGCCATCACCTAGCTGAGATTGAT 63043
QY 221 tLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrrLeuPr 241
Db 63042 GCTTGCTGTCTTTTCATCCCTTCCCTTGTCTTTCTTCTCCACGCTCTTCTTGCC 62983
QY 241 oGluSerProArgTrpLeuValSerIlysglyArgMetAlaGluAlaIlyslsValleuGl 261
Db 62982 CGAGTCCCCAAGGTGGCTGCTGAGCAAAAGGTCGAATGCTTGAAAGCAAAAGCGGTTCTTCA 62923
QY 261 nLySLeuArgGlyLyAspAspValSer----- 270
Db 62922 GAGACTGCGTGGTCCGGAAGATGTGTCTGTTAGTCTTAATCGTCTGACTCTAATTCT 62863
QY 271 -----GlyGluLeuS 274
Db 62862 TTCAATGAACCAATCTTTTCTATTGAATGTTCTTATGTGCAATGTATAGTGAGATGG 62803
QY 274 erLeuLeuLeuGlyLeuGluValIglyIyAspThrSerileGluIuTyrrIleIeg 294
Db 62802 CTTTGTGTTGAGGGTCTTGAATTGAGAGTGAAACAACATAGAGAATATATAATTG 62743
QY 294 lYProAlaThrGluAlaIaAspAspLeuValThraSpGlyAspIySgluGlnIleThrl 314
Db 62742 GTCCCGCGGATGAAAGTACTGATCATGATATAGCTGTGATTAAGATCAAAATTAAGT 62683
QY 314 euTyrrGlyProGluGlyGlnSerTrpIleAlaArgProSerIySglYProIleMetL 334
Db 62682 TATATGTGCAGAAAGGCTGAGTGGGTTGAGGCCAGTCAAAAGGA----- 62633
QY 334 euGlySerValLeuSerIleuAlaSerArgHisGlySer---MetValAsnGlnSerValP 353
Db 62632 --GGAAGCACTATGAGTGTGTTGCTCGCCATGGAAGTACAATAGACAGAGGCAAGGCT 62575
QY 353 roLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaG 373
Db 62574 CATTGATTGATCCTCTGTCACTGTTGGAGCGTTCACGAGAAGATGCCGCACT- 62516
QY 373 lYgIySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThraSpG 393
Db 62515 --GGAAGCATGAGAGTGCCTTGTTCACCATTTTGGAGTATGTTCAAGTGTGGAGGGA 62458
QY 393 lnhIAlaIySAsnGluGlnTrpAspGlyGluAsnLeuHisArgAspAspGluIuTyra 413
Db 62457 ATCAACCAAGACATGAAGATTGGAGTGAAGAAATCTGTTGGAAGAGGTGAGATTATC 62398
QY 413 lAsErAspGlyAlaGlyIyAspTyrrGluAspAsnLeuHisSerProLeuLeuSerArgG 433
Db 62397 CATCCGAC--CATGAGATGATTTCTGAAGATGATCTTCCGTTGATCTCACGTC 62341
QY 433 lnaIaThrGlyAlaGlyIySAspIleValHisHisGlyHisArgGlySerAlaLeuS 453
Db 62340 AAAGCAACAAGCATGAG--AAAGACATGCTCACACTGCTCATGAACTCTTTCTACCT 62284

QY 453 erMetArgArgInThrLeuLeuGlyGluGlyIyAspGlyValSerSerThrAspIleG 473
Db 62283 TCAGACATGGAAGTCAAGTGCAGGGAGCTCAAGGGGAAGAGCGGGTATGATGGGATTG 62224
QY 473 lYgIyGlyTyrrPglNLeuAlaTrpIySTrpSerGluIySgluGlyGluAsnGlyArgIySg 493
Db 62223 GAGGTGATGGCAAGTGGCATGGAATGACGAAAGAGAAGATGAATCGGACAGAAAG 62164
QY 493 lYgIyGlyPheIySArgValTyrrLeuHisGlnGluGlyValProGlySerArgArgIyS 513
Db 62163 AAGGTGGTTTAAACGATATACTTGCATCAAGAAAGTTTCCAGGATCTCGACGTGCT 62104
QY 513 erIleValSerLeuProGlyGlyIyAspValPheGluGlySerGluPheValHisAlaA 533
Db 62103 CAATTGTTTCAATTGCCCTGGTGTGATGGAACCGGTGAG--GCAGATTGTGACAAGCGT 62047
QY 533 lAlaLeuValSerGlnSerAlaLeuPheSerIySglYleuAlaGluProArgMetSera 553
Db 62046 CTGCTTGGTTAGCCCAACCAAGCTCTTATTTCCAAGACCTTCTCAAGAACATACAATTG 61987
QY 553 sPAlaAlaMetValHisProSerGlyValAlaAlaIySglYSerArgTrpIySAspLeuP 573
Db 61986 GTCCTGTATGATGATCATCCGAA--ACAACATAAGGGTCAATTGGCATGATCTTC 61930
QY 573 heGluProGlyValArgArgAlaLeuLeuValGlyValIlegIlnIleleuGln---- 591
Db 61929 ATGATCCTGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAATTACTTCAGCAGG 61870
QY 591 ----- 591
Db 61869 TAGCTAGATAGCCACATCAAAATTAGTAATCACAGAAGAAATAATGCGGTTATGCTA 61810
QY 592 -----GlnPheAlaGlyIleAsnGlyValLeuT 601
Db 61809 ATTCTATGTTTCTGCTTCTCCTTGGAAATGCAGTTCTCAGGCATCAACGAGTCTTT 61750
QY 601 yrrTyrrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerIySPhGlyL 621
Db 61749 ACTACACACCGCAATCTTGAGCAGAGCGGCTGCGGATCTCTACTATCGAACAATGGGGA 61690
QY 621 euSerSerAlaSerAlaSerileuLeuIleSerSerLeuThrThrLeuLeuMetLeuProC 641
Db 61689 TTAGTCTTCTTCAGCATCCTTACTTATAGTCATTGACAACAACCTTGTGATGTACCTG 61630
QY 641 ySileGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
Db 61629 CAATAGCTGTTCATGAGGCTCATGATCTTCTGTCGAAGGTGATTAATACTTAAT 61570
QY 654 ----- 654
Db 61569 TTCCTTGTCTTACAATACACATTAACGAACAAGCTAAATGCCGTTATGAATGTCTT 61510
QY 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeu 669
Db 61509 TTATTTATATTGCAAGACCTTGCTTCTTCACCAAGATACCAATCTGATAGCATCTCTA 61450
QY 670 ValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSer 689
Db 61449 TTGTTTATGTAATCTCAAAATCTTGTTCACATGAACAGCATTTGTCAGCGGCTTATCA 61390
QY 690 ThrValSerValIleValTyrrPheCysCysPheValMetGlyPheGlyProIleProAsn 709
Db 61389 ACCGTAAGCGTGTGCTACTTCTGCTTCTCGTATGGTTCGGTCTGCTCCAAC 61330
QY 710 lIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAla 729
Db 61329 ATCCTCTGTTCAAGATTTTTCCACTCGAGTCCGCGGAATCTGCATCGCATCTGCGCA 61270
QY 730 pheThrPheTrpIleGlyAspIleIleValThrTyrrSerLeuProValMetLeuAsnAla 749
Db 61269 CTCACCTTCTGATCTGTGACATATATCGTCACTTAAGTCTCCCGCTGCTCAAAATCC 61210

QY	750	ILEGLYLEUVALGLYVALPHESERIETRYALAVAlVALCYSLLEULESERPHEVALPHE	769
Db	61209	ATTGAGCTAGCTGTTGTTGGAATGATGACGCAATCGTATGTTGCATTTTCATGGGTCTTT	61150
QY	770	VALPHEULEYSLVALPROGLUThrLYSGLYMETPROLEUGLUVALILEThrGLUPHEPHE	789
Db	61149	GTGTTTCATTAAAGTCCCGAAACTAAAGGCATGCGCACTGCAAGTCATCACAGAGTCTTT	61090
QY	790	ALAVALGLYALALYSGLINALAALA 798	
Db	61089	TCTGTTGAGCTAGACACAGCTGAAGCT 61063	
RESULT 13			
AC136843			
LOCUS	AC136843	135406 bp	DNA linear PLN 06-JUN-2003
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 11 clone		
ACCESSION	OSUNBa0094M06	complete sequence.	
VERSION	AC136843		
KEYWORDS	AC136843.4	GI:31442544	
SOURCE	HTG		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Oryza sativa (japonica cultivar-group)		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1	(bases 1 to 135406)	
AUTHORS	Bhargava,A., Dalal,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M., Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K., Mohapatra,T. and Singh,N.K.		
TITLE	Complete sequence for Oryza sativa chromosome 11 Clone OSUNBa0094M06		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 135406)	
AUTHORS	Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshabu,K., Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bhargava,A., Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-NOV-2002) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India		
REFERENCE	3	(bases 1 to 135406)	
AUTHORS	Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshabu,K., Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bhargava,A., Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-FEB-2003) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India		
REFERENCE	4	(bases 1 to 135406)	
AUTHORS	Bhargava,A., Dalal,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M., Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K., Mohapatra,T. and Singh,N.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JUN-2003) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India		
COMMENT	On Jun 6, 2003 this sequence version replaced gi:28604240.		
FEATURES	Location/Qualifiers		
source	1.135406		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:39947"		
	/chromosome="11"		
	/clone="OSUNBa0094M06"		
BASE COUNT	39875 a	28472 c	27393 g 39665 t 1 others
ORIGIN			
Alignment Scores:			
Pred. No.:	3.15e-146	length:	135406
Score:	2368.00	Matches:	492
Percent Similarity:	73.22%	Conservative:	104

Best Local Similarity: 60.44%			
Query Match: 58.11%			
DB:	8	Mismatches:	156
		Indels:	62
		Gaps:	20
US-10-051-909-32 (1-800) x AC136843 (1-135406)			
QY	19	ARGARGLUARGLEU-----LEUPROSERVALVALLEUVALLEUPROGLYPRO	34
Db	99031	CGACGCACACGGGTGCATGAACCCAAACACCCCGCTGTGTGTCACAAATCATCCA	99090
QY	35	LEUPROPROALA-----SERCYSERSESERGLINGLUPROVALThrSERASP	49
Db	99091	ATCCCTCTCCGCTTTTTCATGTGTTCTCTGAGAAACCGATCCGAGATC-----	99141
QY	50	ASPILEULEUGLUSPLYSMETSERGLYALAVALLEUVALALALEVALASERILEGLY	69
Db	99142	GATCGATCGCCACGTCATGATGAGAGCGCGCTGTGTGGCCGCGCGCCATCGGC	99201
QY	70	ASNLEULEUGLNGLYTRPASPASNALATHRILEALALAVALLEUThryrILEYSLYS	89
Db	99202	AACTACCTGCAGGAGATGGACCAACGCCACCATCGCGCGCTGTGATCAATCAAGCGG	99261
QY	90	GLUPHEGLNLEUGLNASNGLUPROThrVALGLUGLYLEULEVALSERMETSERLEULE	109
Db	99262	GAGTTCGCCCTCGAGACCCAGCCCGCGTGAGGGCTCGTGTGCGCATGTCCCTCATC	99321
QY	110	GLYALATHRILEVALThrThrPHESErGLYPROLEUSERASPserILEGLYARGARGPRO	129
Db	99322	GGCGCCACCATCATCAACCACTTCTCCGCGCCGCTCTCCGACCTCGTCGCGCGCCGCC	99381
QY	130	METLEULEUSERSERILELEUThryrPHEPHESErGLYLEULEILEMETLEUThrSERPRO	149
Db	99382	ATGCTCATCGCT	99441
QY	150	ASNVALThryrVALLEULEULEVALARGPHEVALASPGLYPHEGLYILEGLYLEUVAL	169
Db	99442	AACGTCTACGTCT	99501
QY	170	ThrLEUVALPROLEUThryrLESERGLULILEALPROSERGLULIARGGLYLEULEUASN	189
Db	99502	ACCCTGCTCCCGTCTACATCTCCGAGACCTCCCGCGGAGATCCGCGCGCTCAAC	99561
QY	190	ThrLEUPROGLNPHESERGLYSERGLYMETPHELEUSERThryrCYMETVALPHEGLY	209
Db	99562	ACGCTGCCGACATTCACCGGCTCCGCGGATGTTCATGTCTTACTGATGATCTTCCGC	99621
QY	210	METSERLEUSERPROSERPROASPTRPARGILEMETLEUGLYVALLEUALILEPROSER	229
Db	99622	ATGACGCTCTCGCT	99681
QY	230	LEUPHEPHEPHEGLYLEUThrILEPHEThryrLEUPROGLUSERPROARGThrLEUVALSER	249
Db	99682	CTGCTTACTCTTCT	99741
QY	250	LYSGLYARGMETALAGLUALYSLYSLYVALLEUGLNLYSLEUARGGLYLYSASPVAL	269
Db	99742	AAGGGCGCATGAGAGAGCCAGGCTTGTCTCGAGATGCTGCGCGCGCGAGACGTC	99801
QY	270	SERGLYGLUSERLEULEULEUGLUGLYLEUGLUVALGLYLYASPThrSERILEGLU	289
Db	99802	TCCGCGAGATGGCGCTGCTCTCGAGGGCTCGGACCGCGCGGACACCGAGATCGAG	99861
QY	290	GLUTYrILEILEGLYPROALATHrGLUALALASPSPLEUVALThrASPGLYASPLYS	309
Db	99862	GACTACGTCTGTCGCGCTCTCGAGGGCGAGCC-----GCCGAGAAC	99903
QY	310	GLUGLN-----ILEThrLEUThryrGLYPROGLUGLYGLNserThryrILEALA	325
Db	99904	GAGCAGCGCAGGAGACACCGTCACGCTGTACGGGCGGAGCAGGGGCTTTCGTGGTGCGG	99963
QY	326	ARGPROserLYEGLY-----PROILEMETLEUGLYSERVALLEUSERLEU---ALASER	342
Db	99964	CAGCCGCTGCGCGCGCGCGCGCGGACGATGCTGGGAGAGCTCGTGGGCTGACAGCGCTCG	100023

QY 343 ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe 362
Db 100024 CGGCATGGCAGCATGTACGAGCAG-----ATGAGAGACCCCGGTGGCGCTTTC 100074
QY 363 GlySerValHisGluAsnMetProGlnAlaGly-----GlySerMetArg--- 377
Db 100075 GGGAGCGTCCACGAGCGGCTGCCGAGATCCGGCGCGGCCACCGGACGATGAGGGGG 100134
QY 378 SerThrLeuPheProAsnPhleglySerMetPheSerValThrAspGlnHisAlaLysAsn 397
Db 100135 AGCAGCTGTCTCCCAACCTCGGGAGCATGCTTAGCGTCAACGATAGGCCCGGC--GGC 100191
QY 398 GluGlnTrpAspGluGluAsn-----LeuHisArgAspAspGlu 410
Db 100192 AGCAGCTGGGAGCAGAGAACGTGCAGCTGCGCAGCAGACCTCGACGAGAGAGAG 100251
QY 411 GluTrpAlaSer---AspGlyAlaGlyGlyAspTrpGluAsnLeuHisSerProleu 429
Db 100252 GAGTACCTCTCCGACGACGCGCAAGACGACGACGCGCGCGCGCTCGACGACCGCTG 100311
QY 430 LeuSerArgGlnAlaThrGlyAlaGlyLysAspIleValHisGlyHisArgGly 449
Db 100312 CTGTCCGGCAGAGCACCAGCCTGAGAACCAAGAACGAGCCCGCTCCGGCCAG----- 100365
QY 450 SerAlaLeuSerMetArgArgGlnThrLeuGluGlyGly-----AspGlyVal 467
Db 100366 -----GTGGCATGACGCGCACAGCAGCATCGCGCGCGCGCGCTCGAGACGGCG 100419
QY 468 SerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGly 487
Db 100420 AGCACCATGGGCATCGCGCGGGTGACAGCTGGCGTGAAGTGAGACGAGAACGTGGCC 100479
QY 488 GluAsnGlyArgLysGlyGlyPheLysArgValTyrLeuHisGln-----GluGly 505
Db 100480 CCCGACGGCGCTCAAGCGCGCGGTGAAGCGCATGTACTGACGACGAGAGTCCGAGGCC 100539
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGlu 525
Db 100540 GCGCCCGCGGTGACTCGGGCGCC-----GCCGCGCAGCGC--CAG 100578
QY 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
Db 100579 TCCACGGCGTACGTCCACGCGCGCGCTGTGAGCCGGTGCATGCTGTACACCAAGGAC 100638
QY 546 --LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGlu----- 561
Db 100639 GTCTGTATCGGGCAGAGCCCGACCGACCGCGCTTCGGAACCGCGGAGCGGTGGCG 100698
QY 562 ---ValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla 580
Db 100699 GCGGCGCGCTCGACGCGCGCGCGGTGGCGGAGCTGCTGAGCGCGCGCGCTCCGCCACGCC 100758
QY 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleAsnGlyValLeu 600
Db 100759 CTCTTCTGCGCGCTCACCATCTCATTCAGCAATTCTCCGSCATCAACGCGCTGCTC 100818
QY 601 TyrTyrThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
Db 100819 TACTACACCCCGCAGATCTCGACAGGCGCGCGTCAAGCTCTCTCGCCAGCCTCGGC 100878
QY 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
Db 100879 CTCTCCGCGCACTCCACCTCATCTCATCAGCGGCTTACCAACGCTCTCATGTCTCCG 100938
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Db 100939 TCCATCGCGGTGCGCATGCGCTCATGAGCGCTCGGCGCGCGCGCTCTCTCTG 100998
QY 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
Db 100999 ACGCTGCCGTCTCTGTGCGCTCTCTCGCGCTGCTGTGTGGGGAACGTGTGCCCATG 101058

QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 101059 GCGGCGACCGCGCAGCGCGCGCTGTGCAGCGGGAGCGTCATGCTACTTCTGTGCTTC 101118
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 101119 GTCATGGGTTGCGGCCCATCCCAACATCTCTGCGCGAGATCTTCCGACGAGGGTG 101178
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 101179 AGGGACTCTGCATTCGATCTGCTCGCTGACGCTTGTGCTCGGCGACATCGCGCTCAGC 101238
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 101239 TACAGCTCCCGCTCATGCTCAGCTCCGTGGGCTCGCCGCGTGTCTTCTTACGCC 101298
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 101299 GCCGTGTGCTGCGTGCCTGCTGTGTGCTGAGCGCTCAAGTGCGCCGAGACCAAGGCCCTC 101358
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
Db 101359 CCGCTCGAGGTATCATGAGTTCTTCAACGTGCGGCCCAAG 101400

RESULT 14
AX412656 2205 bp DNA linear PAT 14-JUN-2002
LOCUS AX412656
DEFINITION Sequence 420 from Patent WO0222675.
ACCESSION AX412656
VERSION AX412656.1 GI:21445114
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
1 Glazebrook, J., Wang, X., Dangl, J.L., Eulgem, T. and Zhu, T.
AUTHORS Plant genes, the expression of which are altered by pathogen
TITLE infection
JOURNAL Patent: WO 0222675-A 420 21-MAR-2002;
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (US) ; Eulgem, Thomas (US)

FEATURES
source location/Qualifiers
1..2205
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 521 a 524 c 521 g 639 t
ORIGIN

Alignment Scores:
Pred. No.: 1.3e-142 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
DB: 6 Gaps: 14

US-10-051-909-32 (1-800) x AX412656 (1-2205)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75
Db 1 ATGAAGGAGCGACTCTGCTGCTCTCGCGCCGACCAATCGCAATTTCTTACAGGATGG 60
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 61 GACCAATGCCACCATGTGTGAGCTATGTTATATCAACAAAGACTTGAATCTA----- 114
QY 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CAAACCTCTGTTCAGAGTCTTGTGTTGTATGTATGATCGGTGCAACGCTCATC 171

QY	115	ThrThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSer	134
Db	172	ACGACTTGCTCAGGACCGATATCTGATGGCTCGCAGACGCCCTCATCTTATCA	231
QY	135	SerIleuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu	154
Db	232	TCAGTTATGTATTTGCTGCGGTTTGAATAATGTGTGTCTCCCAATGTCTATGTTCTG	291
QY	155	LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu	174
Db	292	TGCTTGCTAGGCTTCTTAATGGGTTGGTGCCTCGGCTCGCGGTACACTGTCTCCTGTT	351
QY	175	TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe	194
Db	352	TACATTCTGAACCGCTCCTCCGAGATCAGAGACAGTTAAATACTCTCCCTCAGATT	411
QY	195	SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro	214
Db	412	CTTGCTCTGGTGAATGTTTGTGCACTGTATGTGTTTCACTATGTCTCCCTGAGTGAC	471
QY	215	SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly	234
Db	472	TCCCTAGCTGAGAGCCATGCTCGGTGCTCTCGATCCCTCTCTTATTGTTT	531
QY	235	LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIlyGlyArgMetAla	254
Db	532	CTGACGGTCTTTATTGCCCCGAGTCTCTCTGCTGGCTGTAGTAAGAAGATGGAC	591
QY	255	GlyAlaIlybLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSer	274
Db	592	GAGGCTAAGCGAGTCTTCAACAGTTATGTGGCAGAGAAAGATGTAACGATGAGATGGCT	651
QY	275	LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly	294
Db	652	TTACTAGTTGAAGAGACTAGATATAGGAGAGAAAAACAATGAGAACTCTTAGTAAC	711
QY	295	ProGluThrGluAlaAlaAspAspLeuValThrAspGlyAspIlyGlu--GlnIleThr	313
Db	712	TTGGAGGATCATGAAGGTGATGATACACTGGAACCGTTGATGAGGATGACAAATGCGG	771
QY	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerIlyGlyProIleMet	333
Db	772	CTTTATGGAACCCAGAGATCATCTGTAACCTTGCTAGACCTGTCCAGAACAA-----	825
QY	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	826	---AATAGCTCACTGGGCTACGCTCTGCCACGGAGCTTAGCAAAACCAAGCATGATC	882
QY	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	883	CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAGATGCCAGAAGCAGGC	942
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	943	GGAACAACCTCGAGTGGGATTTTCCCTCATTTCCGGAAGCATGTTCACTACTGCCGAT	1002
QY	394	-----HisAlaLysAsnGluInTrpAsp-----GluGluAsnLeuHisArgAsp	408
Db	1003	GCGCTCACGGTAAACCGGCTCATTTGGGAAAAAGACATAGAGAGCATTACAAACAAGAC	1062
QY	409	AspGluGluTyrAlaSerAspGlyAlaGlyIlyAspTyrGluAsp-----AsnLeu	425
Db	1063	AATGATGACTATGCGACTGATGATGGTGGGTGATGATGATGATCTCGGACAACGATTGG	1122
QY	426	HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis	445
Db	1123	CGTAAGCCCTTAATGTCCGCCAGACCAACAGCATGGAC--AAGATATGATCCACAT	1179
QY	446	GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyIlyAsp	465
Db	1180	CCTACAAGTGAAGCACTTAAAGCATGAGACGACACACAGTACCGTTATGCAA--GGCAAC	1236

QY	LOCUS	DEFINITION	SEQUENCE	2254	FROM	PATENT	WO0216655
QY	AX507559	AX507559	2205 bp	DNA	linear	PAT	27-SEP-2002
Db	1237	GGCGAAAGTAGCATGGCAATTGGTGTGGTGGCATATGGGATATAGATACGAAACGAT	1296				
QY	486	GIUGLIGLUASNGLYARGLYSGLUGLYGYPHELYSARGVALTYRLEUHIISGLNGLUGLY	505				
Db	1297	GAA-----TACAGAGGATTATCTTAAGAAAGATGGA	1329				
QY	506	VALPROGLYSERARGARGGLYSERILEVALSERLEUPROGLIGLYASPVALPHEGLU	525				
Db	1330	GCTGAA--TCTCGCCGTGCTCGATCATCTCTATTCGCCGAGAGTCCGGAT-----GGT	1380				
QY	526	GLYSERGLUPHEVALHIALAALALEUVALSERGLNSERALEUPHESERLYGGLY	545				
Db	1381	GGAGGCAGCTACATTCACGCTTCTGCCCTGTGAAGCAGATCTGTCTTGCTCCTAATCA	1440				
QY	546	LEUALAGLUPROARGMETSERASPALAALAMETVALHISPROSERGLUVALAALALYS	565				
Db	1441	-----GTTTCATGGATCCGCCATGGTTCCTCCCGAGAAAATTGCTGCTCT	1485				
QY	566	GLYSERARGTRPLYSASPLEUPHEGLUPROGLYVALARGARGALALEUVALGLYVAL	585				
Db	1486	GGACCACTCTGGTCTGCTCTTCTGAACCTGGTGTAAAGCGTGCTTGGTGTGGTGTCT	1545				
QY	586	GLYILEGLNILELEUGLNGINPHEALAGLYILEASNGLYVALLEUITYRTHRPROGLN	605				
Db	1546	GGCATTCAATACTGCAAGCAGATTTTCAGGTATCAATGAGATTCTCTACTACACTCCTCAG	1605				
QY	606	ILELEUGLUNALAGLYVALAVALILEUSERYSPHEGLYLEUSERSERIALSER	625				
Db	1606	ATTCTCGAAGCGGCTGGCGTAGATATCTCTTCTTCGAGCCTCGACTAAGTTCATCTCT	1665				
QY	626	ALASERILELEUILESERLEUTHRTHRLEUUMETLEUPROCYSLIEGLYPHEALA	645				
Db	1666	GGTCATTCTCCTCATCAGCGGTTTAAACACATTACTCATGCTCCAGCCATTGTCGTTGCC	1725				
QY	646	METLEUUMETASPLEUSERGLYARGARGPHELEULEUGLYTHRILEPROLIELEU	665				
Db	1726	ATGAGACTCATGATGATATCCGGAAGAGGTCAATTCTCTCGACAATCCAGTTCTC	1785				
QY	666	ILEALASERLEUVALILEUVALVALSERASNLEUIEASPLEUGLYTHRILEUALHIS	685				
Db	1786	ATTGTCTCACTGTGCTGCTTGTGCATCAGCAGCTCATCCACATCAGCAAGTCGTGAC	1845				
QY	686	ALALEUSERTHRVALSERVALILEVALTYRPHECYSPHEVALMETGLYPHEGLY	705				
Db	1846	GCAGCACTCTCCACAGGTTGTGCTGCTCTACTCTCTCTCTCGATGGGTACGGT	1905				
QY	706	PROLIEPROASNILEUCYSALAGLUILEPHEPROTHRRARGVALARGGLYLEUCYALIE	725				
Db	1906	CCCATTCCAAACATCCTCTGTCTGTAATCTTCCCAACAAGAGTCCGTGCTCTGCATC	1965				
QY	726	ALALIECYSALAPHETHRPHETRIPILEGLYASPILEILEVALTHRYSERLEUPROVAL	745				
Db	1966	GCCATATGTGCTATGGTCTTTTGGATTGGAGACATTATTGTCACGFACTCATTTCCCGTT	2025				
QY	746	METLEUASNALALIEGLYLEUVALAGLYVALPHESERILETYRALAVALVALCYSEULIE	765				
Db	2026	CTCCTCAGCTCGATCGACTAGTTGGTGTGTTTTCAGCATTTACGCTGCGGTTTGGCTTATC	2085				
QY	766	SERPHEVALPHEVALPHELEULYSVALPROGLUTHIRLYSGLYMETPROLEUGLUVALLIE	785				
Db	2086	TCATGATCTTGGTTTACATGAAAGTCCCGAGACTAAAGGCATGCTTTGGAAGTTATC	2145				
QY	786	THRGLUPHEPHEALAVAGLYALALYSGLNALAALALYSALA	800				
Db	2146	ACAGACTACTTTCCTTTGGAGCT--CAAGCTCAAGCTTCTGCT	2187				

RESULT	15
AX507559	
LOCUS	AX507559
DEFINITION	Sequence 2254 from Patent WO0216655.
ACCESSION	AX507559
	2205 bp DNA linear PAT 27-SEP-2002

VERSION AX507559.1 GI:23388796
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Harper, J.F., Krebs, J., Wang, X. and Zhu, T.
AUTHORS
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 2254 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source 1. location/Qualifiers
1. .2205
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 521 a 524 c 521 g 639 t
ORIGIN
Alignment Scores:
Pred. No.: 1.3e-142 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
DB: 6 Gaps: 14
US-10-051-909-32 (1-800) x AX507559 (1-2205)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db 1 ATGAGGAGGAGGACTCTCGTTGCTCTCGCCGACCAATCGCAATTCTTACAGGATGG 60
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 61 GACAAATGCCACCATGCTGAGCATGTTATATCAACAAGAAGCTGAATCTA----- 114
QY 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CCAACCTCTGTTCAAGGCTTGTCGTTGCTATGTCATGATCGGTGCAACGTCATC 171
QY 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134
Db 172 ACGACTTGCTCAGGACCGATATCTGATGGCTCGGACAGACGCCCATGCTCATTTTATCA 231
QY 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 154
Db 232 TCAGTTATGATTTTCTGTCGCGGTTGATTAATGTTGTGTCCTCCAATGCTATGTTCTG 291
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 292 TGCCTTGCTAGGCTTCTTAATGAGTTGTGGTCCGGCTCGCGGTACACTTGTCCCTGTT 351
QY 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 352 TACATTTCTGAACCGCTCTCCGAGATCAGAGACAGTTAATACTCTCCCTCAGTTT 411
QY 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 412 CTGGGCTCTGGTGAATGTTTGTTCATACTGATGTTTCACTATGTCCTCAGTGAC 471
QY 215 SerProAspTyrPArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db 472 TCCCTTAGCTGAGAGCCATGCTCGGTCTCTCTCGATCCCTTCTTCTTTATTTGTTT 531
QY 235 LeuThrIlePheTyrLeuProGlnSerProArgTyrPLeuValSerLysGlyArgMetAla 254
Db 532 CTCACGGTGTATTATTGCCCCGAGTCTCTCGTTGGCTGTTAGTAAAGAAATGAC 591
QY 255 GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSer 274

Db 592 GAGGCTAAGCGAGTCTTCAACAGTTATGTGGCAGAGAAGATGTTACCGATGATGGCT 651
QY 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly 294
Db 652 TTACTAGTTGAGGACTACATATAGAGAGAGAAAACAATGGAAGATCTCTTAGTAAC 711
QY 295 ProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlu--GlnIleThr 313
Db 712 TTGAGGATCATGAGGTGATGATACACTTGAACCGTTGATGAGATGAGCAATGCGG 771
QY 314 LeuTyrGlyProGluGluGlnSerTyrPleAlaArgProSerLysGlyProIleMet 333
Db 772 CTTATGGAACCCACGAAATCAATCGTACCTTGCTAGACCTGCCAGAACAA----- 825
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 826 ---AATAGCTCACTGGGCTACGCTCTCGCCACGGAAGCTTAGCAACCAAGCATGATC 882
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 883 CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAAGATGCCAGACGAGC 942
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 943 GGAACAACCTCGAGTGGGATTTCCCTCATTTCCGAAGCATGTTCAGTACTACTGCCGAT 1002
QY 394 -----HisAlaLysAsnGluGlnTyrAsp-----GluGluAsnLeuHisArgAsp 408
Db 1003 GCGCCTCACGTTAAACCGGCTCATTTGGAAAGACATAGAGACCATTACAAACAAGAC 1062
QY 409 AspGluGluTyrAlaSerAspGlyAlaGlyLysAspTyrGluAsp-----AsnLeu 425
Db 1063 AATGATGACTATGCGACTGATGATGTGCGGGTGAATGATGACTCGACAAACGATTTG 1122
QY 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
Db 1123 CGTAGCCCTTAATGTCGCCGCCACAGACCAAGCATGAC--AAGATATGATCCACAT 1179
QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGluGlyGlyAsp 465
Db 1180 CCTACAAGTGGAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCAA--GGCAAC 1236
QY 466 GlyValSerSerThrAspIleGlyGlyIleTyrGlnLeuAlaTyrLysSerGluLys 485
Db 1237 GGGCAAAAGTAGATGGAATGTGTGTGGTGGCATATGGGATATAGATACGAAACGAT 1296
QY 486 GluGlyGluAsnGlyArgGlyGluGlyPheLysArgValTyrLeuHisGlnGluGly 505
Db 1297 GAA-----TACAAGAGTATATCTTAAGAAGATGGA 1329
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyLysAspValPheGlu 525
Db 1330 GCTGAA---TTCGCCGTGCTGATCATCTCTATTTCCCGAGAGTCCGAT-----GGT 1380
QY 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
Db 1381 GGAAGCAGCTACATTACGCTTCTGCCCCCTGTAAGCAGATCTGTTCTTGCTCTAAATCA 1440
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
Db 1441 -----GTTATGGAATCCGCCATGTTCCCGGAGAAATGTGCTCT 1485
QY 566 GlySerArgTyrLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal 585
Db 1486 GGAACAACCTGCTGCTGCTTCTTGAACCTGCTGAACGCTGCTGTTGTTGGTGTG 1545
QY 586 GlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
Db 1546 GGCATTCAATACGACGAGCTTTTCAGGTATCAATGAGTCTCTACTACACTCTCTCAG 1605
QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625


```
Db 1606 ATCTCGAAGCGGCTGGCGTAGATATTCTTCTTCGAGCCTCGACTAAGTTCATCTCT 1665
QY 626 AlaserlleuileuilesexserleuthrthrleuileumetleuproCysileglypheAla 645
Db 1666 GCGTCATTCCTCATCGCGTTTAACAACATTACTCATGCTCCACGCCATTGTGCTGCC 1725
QY 646 MetleuleumetAspleuSerclyArgArgpheleuleuGlyThrileProileleu 665
Db 1726 ATGAGACTCATGGATGATATCCGAAGAAGTCATTACTTCTCTGAGACAATCCCAATTCTC 1785
QY 666 lleAlaserleuValilleuValValSerAsnleuileaspleuGlyThrleuAlaHis 685
Db 1786 ATTGTCTCACTGTGCTGCTCTTCATCAGCGAGCTCATCCACATCAGCAAGTCGTGAAC 1845
QY 686 AlaileuSerThrValSerValileValTyrpheCyspheValmetGlypheGly 705
Db 1846 GCAGCACTCTCCACAGGTTGTGTCTGCTCTACTTCGTCTCTGATGGGTTACGGT 1905
QY 706 ProileProAsnilleuCysAlaGluilepheProthrArgValArgGlyLeuCysile 725
Db 1906 CCCATTCCAAACATCCTGTGTCTGAATCTTCCCAACAGAGTCCGTGCTCTGCATC 1965
QY 726 AlaileCysAlaphethrPheTrpileGlyAspileileValThrTyrSerleuProVal 745
Db 1966 GCCATATGTGCTATGCTTTTGGATTGAGATGAGACATTATGTACAGTACTCACTCCCGTT 2025
QY 746 MetleuAsnAlaileGlyleuAlaGlyValPheSerileTyrAlaValValCysleuile 765
Db 2026 CTCCTCAGCTCGATCGGACTAGTTGGTGTTCAGCATTTACGCTGCGGTTTGCCTTATC 2085
QY 766 SerpheValpheValpheleuLysValProGluThrlysglyMetProleuGluValile 785
Db 2086 TCATGGATCTTCGTTTACATGAAGAATCCCGAGACTAAAGGCATGCCCTTGGAGTTATC 2145
QY 786 ThrGluPhepheAlaValGlyAlaLysGlnAlaAlaAlaLysAla 800
Db 2146 ACAGACTACTTTGCCCTTTGGAGCT--CAAGCTCAAGCTTCTGCT 2187
```

Search completed: January 5, 2004, 21:44:56
Job time : 8849 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:39:20 ; Search time 41 Seconds
(without alignments)
1876.462 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWAVQTPFPDLDLR.....PLEVITTEFFAVGAKQAAXA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2584	63.4	729	2	T06127	probable sugar tra
2	2278	55.9	734	2	H86340	Sugar transporter
3	2245	55.1	729	2	T45780	sugar transporter-
4	2183	53.6	734	2	T51139	sugar transport pr
5	587.5	14.4	457	2	E70070	metabolite transpo
6	536	13.2	580	2	D66426	hypothetical prote
7	533	13.1	461	2	D70073	metabolite transpo
8	524.5	12.9	580	2	D84772	probable sugar tra
9	520.5	12.8	582	2	F71431	hypothetical prote
10	518.5	12.7	469	2	H97064	probable sugar-pro
11	505	12.4	473	2	G69789	sugar transporter
12	493.5	12.1	508	2	G84564	probable sugar tra
13	493	12.1	493	2	A85433	sugar transporter
14	489	12.0	521	2	G84864	probable membrane
15	488	12.0	549	2	T14606	probable sugar tra
16	486.5	11.9	482	2	B69803	metabolite transpo
17	474	11.6	547	2	C84593	probable sugar tra
18	473.5	11.6	464	2	F69587	L-arabinose transp
19	465	11.4	419	2	E69888	metabolite transpo
20	462	11.3	612	2	B40538	myo-inositol trans
21	460.5	11.3	511	2	A84537	probable sugar tra
22	459.5	11.3	511	2	H84536	probable sugar tra
23	458	11.2	560	2	T51485	sugar transporter-
24	451	11.1	464	2	F65079	galactose-proton s
25	448	11.0	464	2	C91106	galactose-proton s
26	445	10.9	464	2	F85951	galactose-proton s
27	435.5	10.7	471	2	AB0868	L-arabinose isomer
28	434.5	10.7	472	2	S47089	arabinose-proton s

30	434	10.7	491	2	A26430	xylose transport p
31	434	10.7	491	2	F91255	xylose-proton symp
32	434	10.7	491	2	B86096	xylose-proton symp
33	432.5	10.6	490	2	T14545	probable sugar tra
34	432	10.6	557	2	T38125	myo-inositol trans
35	431	10.6	547	2	A48442	membrane transport
36	429.5	10.5	472	2	B26430	L-arabinose isomer
37	429.5	10.5	472	2	B91091	L-arabinose isomer
38	429.5	10.5	472	2	E85936	L-arabinose isomer
39	421.5	10.3	584	2	S69555	myo-inositol trans
40	418.5	10.3	468	2	S10014	glucose transport
41	417.5	10.2	555	2	S69671	hypothetical prote
42	415	10.2	606	2	T27072	hypothetical prote
43	412	10.1	487	2	E96782	hypothetical prote
44	412	10.1	502	2	B70845	probable sugar tra
45	407.5	10.0	613	2	T27077	hypothetical prote

ALIGNMENTS

RESULT 1
T06127
probable sugar transport protein F23E12.140 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
C;Accession: T06127
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Hohnsels,
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15485
A;Accession: T06127
A;Molecule type: DNA
A;Residues: 1-729 <BEV>
A;Cross-references: EMBL:AL022604; GSPDB:GN00062; ATSP:F23E12.140
C;Genetic: A;Experimental source: cultivar Columbia; BAC clone F23E12
A;Gene: ATSP:F23E12.140
A;Map position: 4
A;Introns: 27/1; 216/1; 433/1; 519/3; 582/2
C;Keywords: sugar transport; transport protein

Query Match	Score	2584	DB 2	Length	729
Best Local Similarity	68.8%			Pred. NO. 9.7e-174;	
Matches	512;	Conservative	89;	Mismatches	123;
				Indels	20;
				Gaps	8;
QY	56	MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQLONEPTVEGLIVMSLIGATIVT	115		
Db	1	MSGAVLVAIAAAGVNLQGMNATIAGAVLYIKKEFNLESNPVEGLIVAMSLIGATLIT	60		
QY	116	TFSGPLSDSIGRRPMLILSSILYFSGGLIMSPVYVLLARFVDGFGIGLAVTLVPLY	175		
Db	61	TCSGGVADWLGRPMILSSILYFVGSVLMMSPNVYVLLGRLLDFGVGLVTLVPIY	120		
QY	176	ISEIAPSEIRGLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMGVLAIPSLFFGL	235		
Db	121	ISEIAPPEIRGLNTLPQFTGSGGMFLSYCMVFGMSLSPSPWRMLGVLFIPSLVFFFL	180		
QY	236	TIFYLPESPRWLVSQGRMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSTIEEYIGP	295		
Db	181	TVEFLPESPRWLVSQGRMLEAKRVLOQLRGREDVSGEMALLVEGLIGETTIEEYIIGP	240		
QY	296	ATEAADLVTGDKKQITLYGPEEGQSWIARPSKPIMLGSVLASRHGS-MVNQSVPL	354		
Db	241	ADEVTTDDHDIAVDKQIKLYGAEBGLSWVARPVKG---GSTMSVLSRHGSTMSRRQGS	296		
QY	355	MDPIVTLFGSVHENMQAGGSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEYAS	414		
Db	297	IDPLVTLFGSVHEKMDPT-GSMRSALFPHFGSMFSGVGNQPRHEDWDEENLVGEGEDYPS	355		
QY	415	DGAGGDYEDNLHSPLSQATGAEGKDIVHGHRSALSMRQTLGEGGDVYSTDIDG	474		
Db	356	D-HGDDSEDDLHSPLSISQRTTSMK-KDMPTAHGTLSTFRHGSQVQAGGAGGSMGIGG	413		

Db 181 AAFPLESPRWLVSKGRMDARQVLRGRREDVSGELALLVEGLGVGKDTISIEEYVIGP 240

QY 296 ATEAADLVTDGKEQITLYGPEEGSWIARPSKGPIMLGSVLSLARHGSMVNOQVPLM 355

Db 241 DNEBEGNELPRKDQIKLYGEBEDQOSWMAKPVKGQ---SSLALASROGSMLEPRGSLM 296

QY 356 DPLVTLFGSVHENMP--QAGGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDDEYA 413

Db 297 DPLVTLFGSIHENLPSENNANASRSMLFPNMGSLGMGR--QESQWDP---RNMD-- 349

QY 414 SDGAGDYEDNLHSPILSRQATGAEGKDIVHGHGRGSLSMRQTL-LGEGGDVSTDI 472

Db 350 ----SSQDENLNSPLSPQT--EPDD--YHQRVTGTHRRQSSLFMANVGETATATSI 401

QY 473 GGGWQLAMKMKSEKEGENGRK-EGGFKRVYLHOE-----GVPGSRGSIIVSL-PGGGDV 523

Db 402 GGGWQLAMKYNKDKVGADGKRVNGQLQRMVTHEETANNNTNNIPFSRRGSLSFHPEDGH 461

QY 524 FEGSEFVHAALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRRALLV 583

Db 462 DQVNGYVQAALVVSQASMMPGKGETAML-----PKEV-KDGPGRRELKEPGVKRALMV 514

QY 584 GVGIOILQOFAGINGVLYTTPQILEQAGVAVILSKFGLSASASILISLTLLMLPCIG 643

Db 515 GVGIOILQOFAGINGVMYTTPQILEETGVSSLLTNLGISASASALLISALTLLMLPCI- 573

QY 644 FAMLMDLSGRFLLGTPIPIIASLVILVNSNLDLGLAALLSTVSVTVYFCCFVMG 703

Db 574 -----LVMSRSLMLSTIPILISLTVLVIGSLVNLGGSINALISTASVTVYLSCFVMG 626

QY 704 FGPPIPLCAEIPFTRVRGLCIAICFTFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVC 763

Db 627 FGAIPNLCSEIPTSVRGCLTITICALTFCWICDIIVTYPMLKSIGIAGVFGIYAVIC 686

QY 764 LISFVEFLKVPETKGMPLLEVITEFFPAVGAQAQAAKA 800

Db 687 AVAMVEVYLKVPETKGMPLLEVISEFPFVGAKQDAAA 723

RESULT 4

T51139

sugar transport protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000

C;Accession: T51139

R;Tjaden, J.; Neuhaus, E.

submitted to the EMBL Data Library, August 1995

A;Description: A new sugar transport protein from Arabidopsis thaliana.

A;Reference number: Z25311

A;Accession: T51139

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-734 <TJA>

A;Cross-references: EMBL:Z50752; PIDN:CAA90628.1

Query Match 53.6%; Score 2183; DB 2; Length 734;

Best Local Similarity 59.5%; Pred. No. 1.6e-145;

Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;

QY 56 MSGAVLVAIVASIGNLQGWMDNATIAAALYIKKEFQLQNEPT-VEGLIVMSLIGATIV 114

Db 1 MKGATLVALAATIGNFLQGWMDNATAGAMVYINKDLNL---PTSVQGLVAMSLIGATVI 57

QY 115 TTFSGPLSDSIGRRPMLILSSILYFSGLIMLMSPNVYVLLARFVDFGIGLAVTVPL 174

Db 58 TTCGSPISDMLGRPMLILSSVMYFVCGLIMLMSPNVYVLCFARLLNGFGAGLAVTVLPV 117

QY 175 YISEIAPSEIRGLNLTLPQFSSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPSLFFFG 234

Db 118 YISETAPPEIRGQLNTLPQFLSGGMFLSYCMVFTMSLSDSPSWRAMLGVLISPSLLYLF 177

QY 235 LTIFFYLPSPRWLVSKGRMAEAKVYQKLRKDDVSGELSLLEGLLEVGGDTISIEEYIG 294

Db 178 LTVFYLPSPRWLVSKGRMDAKRVLQQLCGREDVTDDEMALLVEGLDIGEKTMEDLLVT 237

QY 295 PATEAADLVTDGKE-QITLYGPEEGSWIARPSKGPIMLGSVLSLARHGSMVNOQVPLM 353

Db 238 LEDHEGDDTLETVEDEGQIRLYGTHENQSYLARPVPEQ---NSSLGRSRHGSLANQSMI 294

QY 354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQ--HAKNEQW--EENLHRD 408

Db 295 LKDPLVNLFGSLHEKMEAGNGTRSGIFPHGSMFSTTADAPHGKPAHWKXDIESHYNKD 354

QY 409 DEEYASDGAGDYED--NLHSPILSRQATGAEGKDIVHGHGRGSLSMRQTLLEGGGD 465

Db 355 NDDYATDDGAGDDDDSDNDLRSPLMSRQTSMD-KDMIPIPTSGSTLSMRHSTLMQ-GN 412

QY 466 GVSSTDIGGWQLAMKMKSEKEGENGRKEGFRVYLHOEGVPGSRGSIIVSLPGGDVFE 525

Db 413 GESSMGIGGGMHGYRYENDE-----YKRYLLKEDGAE-SRRGSIISIPGPD-G 460

QY 526 GSEFVHAALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRRALLVGV 585

Db 461 GGSYTHASALVSRSVLPKPS-----VHGSAMVPPEKTAASGPLMSALBEGVKRALVGV 515

QY 586 GIOILQOFAGINGVLYTTPQILEQAGVAVILSKFGLSASASILISLTLLMLPCIGFA 645

Db 516 GIOILQOFAGINGVLYTTPQILERAGVDILSSLGSSISASFLISGLTLLMLPAIIVA 575

QY 646 MLMDLSGRFLLGTPIPIIASLVILVNSNLDLGLAALLSTVSVTVYFCCFVMGFG 705

Db 576 MRLMDVSGRRSLMLTIPVILVSLVVLVISELIHISKVNNAALSTGCVVLYFCFVVMGYG 635

QY 706 PIPNILCAEIPFTRVRGLCIAICFTFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVCLI 765

Db 636 PFQTSVLKSSQADRGLCIAICAMFWIGDIIVTYSLPVLLSIELVGVSIIYAAVCI 695

QY 766 SFVFEFLKVPETKGMPLLEVITEFFPAVGAQAQAAKA 800

Db 696 SWIFVYMKVPETKGMPLLEVITDYFAFGA-QAQASA 729

RESULT 5

E70070

metabolite transport protein homolog ywtG - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C;Accession: E70070

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrarri, B.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.

koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: E70070

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-457 <KUN>

A;Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PIDN:CAB15600.1; PID:g2636109

A;Experimental source: strain 168

C;Genetics:

A;Gene: ywtG

C;Superfamily: glucose transport protein

Query Match 14.4%; Score 587.5; DB 2; Length 457;

Best Local Similarity 22.8%; Pred. No. 1.4e-33;

Matches 168; Conservative 102; Mismatches 162; Indels 305; Gaps 12;			
QY	55	KMSGAVLVAIVASIGNLQGNATIAAAVLVIKKEFQONEPTVEGLIVSMSLGATIV	114
Db	2	KKQSNIMWLYFFGALGYDPTGVISGAILFMKKEGL--NAFTEGLVSSILVGAILG	59
QY	115	TTFSGPLSDSIGRRPMLISLILYFSGSLIMLWSPNVYVLLARFVDFGIGLAVTLVPL	174
Db	60	SGAAGKLTDRFGRKKAIMAAALLFCIGGLVALAPNTGVAVLFRILGLAVGSTTIVPL	119
QY	175	YISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDPWRIMLGVLAIPSLFFFG	234
Db	120	YLSELAPKHKRGALSSLNQMLTVGILLSY--IVNYIFADAEMAKRWMLGLAAVPSLLL-	176
QY	235	LTIFFLPESPRWLVSCKGRMAEKVLOKLRKDDVSGELSLLEGLVGGDTSEIYYIG	294
Db	177	IGILFMPESPRWLFTNGEESKAKKILEKRGTKDI-----	211
QY	295	PATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVPL	354
Db	212	-----	211
QY	355	MDEIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSYTDQAKNEQWDEENLHRDEEYAS	414
Db	212	-----DQE---	214
QY	415	DGAGGDYEDNLHSPLLSRQATGAEGKDIVHGHGSAISMRRQTLLEGGDVSSSTDIGG	474
Db	215	-----IH-----DI--	218
QY	475	GMQLAMKWKSEKENGKRGKRVYLLHOGVPGSRGISTVSLPGGDVFEGSEFVHAAA	534
Db	219	-----KEAEKQDEGLKE-----	231
QY	535	LVQSQALFSKGLAEPKMSDAAVHPSEVAAGSRWKDLFEPGVRALLVGIGIQIQQFA	594
Db	232	-----LFPDWVRPALIAGLAFLOQFI	254
QY	595	GINGLVYTPQILEQAGVAVILSKFGLSSASIL-----ISSLTLLMPCIGFAMLM	650
Db	255	GTNTIITYAPKFTNVG-----FG--NSASILGTVGIGTVNVLMTL---VAIKIID	300
QY	651	LSGRFLLGTIPILIASLVILVSNLIDGLTAHALSTVSIVVYFCFVMGFGPIPI	710
Db	301	KIGRKPLLLFGNAGMVISLVLALVNLFPNNTPAASMTVLCIGVIVFAVSMGPVAV	360
QY	711	LCAEIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVNLNAIGLAVESIYAVVCLISFEV	770
Db	361	MLPELPLHVRGIGTGVSTLMLHVTLLIVSLTYPIIMEAIGISYFLIYAIGMAFLFV	420
QY	771	FLKVPETKGMPLVITE 787	
Db	421	RFKVTETKGRSLEEIEQ 437	

RESULT 6

D86426
hypothetical protein F12P21.2 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C/Accession: D86426
R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: D86426

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-580 <STO>
A/Cross-references: GB:AE005172; NID:g11120774; PIDN:AAG30955.1; GSPDB:GN00141
A/Map position: 1
C/Superfamily: glucose transport protein

Query Match 13.2%; Score 536; DB 2; Length 580;
Best Local Similarity 21.6%; Pred. No. 8.4e-30;
Matches 175; Conservative 93; Mismatches 169; Indels 372; Gaps 11;

QY	62	VAIVASIGNLQGNATIAAAVLVIKKEFQONEPT-VEGLIVSMSLGATIVTTFSSGP	120
Db	31	LAFSAGIGGLFGYDTGVISGALLYIRDDFKSVDRNTWLQEMIVSMAYAGAIYGAIGGW	90
QY	121	LSDSIGRRPMLISLILYFSGSLIMLWSPNVYVLLARFVDFGIGLAVTLVPLIYSEIA	180
Db	91	ANDKLGRSAIIMADFLFLGAIIMAAAPNPSLLVGRVFGVGLGVMASMTAPLYISEAS	150
QY	181	PSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDPWRIMLGVLAIPSLFFGLTIFYL	240
Db	151	PAKIRGALVSTNGFLITGGQFLSYLINLAFI-DVTGTWRMLGIAGIPALLQFVL-MFTL	208
QY	241	PESPRWLVSCKGRMAEKVLOKLRKDDVSGELSLLEGLVGGDTSEIYYIGPATEAA	300
Db	209	PESPRWLVRKGREBAKAILRIRYSAEDVEGEIRALKDSVET-----	250
QY	301	DDLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVPLMDPIVT	360
Db	251	-EILEEGSSEKINMI-----	264
QY	361	LFGSVHENMPQAGSMRSTLFPNFGSMFSYTDQAKNEQWDEENLHRDEEYASDAGGD	420
Db	265	-----	264
QY	421	YEDNLHSPLLSRQATGAEGKDIVHGHGSAISMRRQTLLEGGDVSSSTDIGGQQLAW	480
Db	265	-----KLCKAKT-----	271
QY	481	KWSEKEGENGRKEGFKRVYLLHOGVPGSRGISTVSLPGGDVFEGSEFVHAAALVSQSA	540
Db	272	-----	271
QY	541	LFSKGLAEPKMSDAAVHPSEVAAGSRWKDLFEPGVRALLVGIGIQIQQFAGINGVL	600
Db	272	-----VRGLIAGVGLOVFGQFVGINTVM	295
QY	601	YTPQILEQAGVAVILSKFGLSSASASILSSLTLLMPCIGFAMLMDSGRFLLG	660
Db	296	YSPITVQLAGFA-----SNRTALLSLVTAGLNAGFSIISYFIDRIGRKULLII	346
QY	661	TIPILIASLVIL-----VSNLIDLT-----	682
Db	347	SLFGVILSLGILTVFEAATHAPAISLETORFNNISCPDYKSAMNTNAMDCTCLKAS	406
QY	683	-----LAH-----ALLSTVSVIVY	696
Db	407	SPSCGYCSSPIGKHPGACWISDSVKDLCHENRMLYTRGCPSNFGWFALLGLGYIIF	466
QY	697	FCCFWMGFGPIPIILCAEIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVNLNAIGLAVF	756
Db	467	FS---PGMGTVPMIVNSEIYPLRFRGICGGIAATANWISNLIVAQSFSLTEAIGTSWF	523
QY	757	SIYAVVCLISFVFVLKVPETKGMPLVITE 785	
Db	524	LIFGVISVIALLFVAVCVPETKGMPEEI 552	

RESULT 7

D70073
metabolite transport protein homolog yxoc - Bacillus subtilis
C/Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: D70073
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D70073
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-461 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16017.1; PID:g2636527
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxcC
C;Superfamily: glucose transport protein

Query Match	13.1%;	Score 533;	DB 2;	Length 461;
Best Local Similarity	21.6%;	Pred. No. 9.9e-30;		
Matches	157;	Conservative 102;	Mismatches 168;	Indels 300; Gaps 12;
QY	61	LVAIVASIGNLQGMNATIAAVLYIKKEPOLQNEPTVEGLIVSMSLIGATIVTTSGP	120	
DB	9	MIYFFGALGGLLYGYDTGVISGALLFINNDIPLTT-LTEGLVVSMLLGAIFGSALSGT	66	
QY	121	LSDSIGRRPMLILSSILYFFSGILMWSPNVYLLARFVDGFGIGLAVTLVPLYISEIA	180	
DB	67	CSDRWGRKRVFVLSTIFIGALCAFSQTIIGMLIASRVILGLAVGSGTALVPVLSSEMA	126	
QY	181	PSEIRGLNLTLPQFSSGGMFLSYCMVFGMSLSPSPDRIMLGVLAIPLFFFGLTIFYL	240	
DB	127	PTKIRGLTGMNMLMTVTGILLAY--IVNYLFTPEAWRMVGLAAPPVALL-IGIAFM	183	
QY	241	PESPRWLVSCKRMAEAKVLOKLRKDDVSGELSLLEGLVGGDTSEIYIIGPATEAA	300	
DB	184	PESPRWLVRKGSSEEARIMNITHDPKDIEMELAEKQG-----	222	
QY	301	DDLVTGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVPLMDPIVT	360	
DB	223	-----EAEKETTL-----	231	
QY	361	LFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDDEYASDAGGD	420	
DB	232	-----	231	
QY	421	YEDNLHSPPLSRQATGAEGKDIVHHGHRGALSMMRQTLLEGGDGVSSTDIGGWQLAW	480	
DB	232	-----	236	
QY	481	KWSEKEGENGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGDIVFEGSEFVHAALVSQSA	540	
DB	237	KW-----	238	
QY	541	LFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALVGVGIQLQOFAGINGVL	600	
DB	239	-----	262	
QY	601	YTTPOILEQAGVAVILSKFGLSSASILISLTLMLPCIGFAMLLMDLSGRRFLTL-	659	
DB	263	YAAP-----	312	
QY	660	GTIPILIASLVILVSNLIDGLTAH-ALLSTVSVIVFCFVWGFGPIPNILCAEIPPT	718	

Db 313 GSVGITLSIALSGV--LTLGLSASTAMTVFLGVYIVFYQATMGCPVWVLMPELFPS 370
QY 719 RVKGLGIAICAFITWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFLKVPETK 778
Db 371 KARGAATGFTTLVISAANLIVSLVPLMLSAMGIAVFMVFSVCLLSFFFAFYMPETK 430
QY 779 GMPLEVI 785
Db 431 GKSLEEI 437
RESULT 8
D84772
probable sugar transporter [imported] - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C;Accession: D84772
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84772
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <STO>
A;Cross-references: GB:AE002093; NID:g4263781; PIDN:AAD15441.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g35740
A;Map position: 2
C;Superfamily: glucose transport protein

Query Match	12.9%;	Score 524.5;	DB 2;	Length 580;
Best Local Similarity	20.9%;	Pred. No. 5.4e-29;		
Matches	171;	Conservative 102;	Mismatches 175;	Indels 369; Gaps 13;
QY	62	VAIVASIGNLQGMNATIAAVLYIKKEF-QLONEPTVEGLIVSMSLIGATIVTTFSGP	120	
DB	29	LALSAGIGLLFGYNTGVIAGALLYIKKEFGEVDNKTWLQETIVSMTVAGAIAGAGW	88	
QY	121	LSDSIGRRPMLILSSILYFFSGILMWSPNVYLLARFVDGFGIGLAVTLVPLYISEIA	180	
DB	89	YNDKFRMSVLADVIFLLGALVMVIAHAPVYIILGRLLVGFVGMASMTSPLYISEMS	148	
QY	181	PSEIRGLNLTLPQFSSGGMFLSYCMVFGMSLSPSPDRIMLGVLAIPLFFFGLTIFYL	240	
DB	149	PARIRGALVSTNGLLITGGQFLSYLINLAFAVHTPG-TWRMWLGVSAPALIQFCL-MLTL	206	
QY	241	PESPRWLVSCKRMAEAKVLOKLRKDDVSGELSLLEGLVGGDTSEIYIIGPATEAA	300	
DB	207	PESPRWLVRNDRKABSRDILERTYPAEMVEAEIAALKESVR-----	251	
QY	301	DDLVTGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSVPLMDPIVT	360	
DB	252	DE-----	253	
QY	361	LFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDDEYASDAGGD	420	
DB	254	-----	253	
QY	421	YEDNLHSPPLSRQATGAEGKDIVHHGHRGALSMMRQTLLEGGDGVSSTDIGGWQLAW	480	
DB	254	-----	258	
QY	481	KWSEKEGENGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGDIVFEGSEFVHAALVSQSA	540	
DB	259	TFSDK-----	267	
QY	541	LFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALVGVGIQLQOFAGINGVL	600	
DB	268	-----	296	

Db	193	-ITLFEVESPRLVKSNGIKKAAAVLTKINGAEIAKQELDISKSLATENDSSL-----	246
QY	294	GPATEAADLVTDGDKQITLYGPEEGSWIARPSKGPIMLGSVLSASRHGSMVNQSV	353
Db	247	-----GQ-----	248
QY	354	LMDFIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDEEYA	413
Db	249	-----	248
QY	414	SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRGSALSMRRQTLLEGGDVSTDIG	473
Db	249	-----	248
QY	474	GGWQLAMKWEKEGENRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDVFEGSEFVHA	533
Db	249	-----	248
QY	534	ALVSQSALFSKGLAEPMSDAMVHPSEVAAKGRKDLFEPGVRALLVGVGIQLQOF	593
Db	249	-----LIQPLRRALLIGIFLAINQA	270
QY	594	AGINGVLYTPQILEQAGVAVILSKFGLSSASASIL--ISSLTLLMPCIGFAMLLMDL	651
Db	271	IGMNSITTYGPEIFQMIGF-----KNSSFLATSVIGVEVFSTIL-----AMFLIDK	318
QY	652	SGRRFL-LGTIPILASVILVSVNLIDGLTAHALSTVSVI-VYFCCFVMGF-----G	705
Db	319	LGRKKLMEIGSAAWAFVLLI-----GTSFYIKLSNGFVILFIICFVVSFCISMG	369
QY	706	PIPNILCAEIPFTRVRGLCIAICFTWIGDIIVTYVSLPVMNAIGLAGVFSIYAVVCLI	765
Db	370	PIPWIMPEIFPNHLRARATGATIFLWGANWAIQGFPMLLNGIGGAYTFWIFCGINVI	429
QY	766	SFVFVFLKVPETKMPLEVITEFFAVGAKQAA	797
Db	430	CFLVVTTKVPETKNKSLTEEIEKFWIPKSKONA	461

RESULT 11
G69789
sugar transporter homolog ydjK - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C/Accession: G69789
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chn
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Toseato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69789
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-473 <KUN>
A/Cross-references: GB:Z99107; GB:AL009126; NID:g2632866; PIDN:CAB12442.1; PID:g2632936
A/Experimental source: strain 168
C/Genetics:
A/Gene: ydjK
C/Superfamily: glucose transport protein
Query Match 12.4%; Score 505; DB 2; Length 473;

Best Local Similarity		22.3%;	Pred. No. 9.6e-28;	
Matches		168;	Conservative 106;	Mismatches 176; Indels 302; Gaps 16;
QY	54	DKMSGAVLVAIVASIGNLLQGWDNATIAAAVL	YIKKEFQLONEPTEGLIVMSLIGATI	113
Db	6	NQMSFLRTIILVSTFGGLIFGYDTGVNLGALPYMGE	PDQLNNAFTGELVTSLLFGAAL	65
QY	114	VTPSGPLSDSIRRRPMLILSSILYFSGLIMLSPNVYVLLARFVDGFIGLAVTLVP		173
Db	66	GAVFGGRMSDFNGRRKNILFLAVIFFISTIGCTFAPNVTWIIISRFVLGI	AVGASVTP	125
QY	174	LYISEIAPSEIRGLNLTLPQFSGSGMFLSYCM--VFQMSLSP-SPDWRIMLGVLAIPSL		230
Db	126	AYLAEMSPVESRGMVTQNELMIVSGQLLAFVFNALIGTTMGDNSHVRFMLVIALPAL		185
QY	231	F-FFGLTFYLPESPRWLVS	SKGRMAEAKVLOKLRKDDVSGELSLLEGLEVGDTSIE	289
Db	186	FLFFGM--IRMESPRWLVS	SKGRKEDALRYLKIR	218
QY	290	EYIIGPATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVN		349
Db	219	----- 218		
QY	350	QSVPLMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRD		409
Db	219	-----D 219		
QY	410	EYASDAGAGDYEDNLHSPILSRQATGAEGKDIVHHGRGSALSMRRQTLLEGGDVSS		469
Db	220	EKRAA-----AEIOEI-----		230
QY	470	TDIGGGWQLAMKWEKEGENRKEGFKRVYLLHQEGVPGSRGSIIVSLPGGDVFEGSEF		529
Db	231	-----EPAFKK----- 236		
QY	530	VHAALVSQSALFSKGLAEPMSDAMVHPSEVAAKGRKDLFEPGVRALLVGVGIQI		589
Db	237	-----EDQLEKAT-----FKDLSVPWVRIRIVFIGLGIAI 265		
QY	590	LQQFAGINGVLYTPQILEQAGVAVILSKFGLSSASASILSSLTLLMPCIGFAMLLM		649
Db	266	VQQITGVNSIMYGTIELRNSG--FQTEAALIGNIANGVISLATFV-----GIWLL		315
QY	650	DLGRR-FLLIGTIPILIASVLIVSVNLIDGLTAHALSTVSVIYVPCCFVMG-FGPI		707
Db	316	GRVGRPRMLMTGLIGTTALLIGIFSLVLE-GSPALPYV-VLSLVTFLAFQGAISPV		373
QY	708	PNILCAEIPFTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISF		767
Db	374	TWMLSEIFPLRLRGLMGVTFCLMNVNFAVSFTFPILLAIGLSTTFFIFVGLGICSV		433
QY	768	VFVFLKVPETKGMPLLEVITEFFAV---GAKQ		795
Db	434	LFVKRFLPETKGLSLEQLEENFRAYDHSGAKK		465

RESULT 12
G84564
probable sugar transporter [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C/Accession: G84564
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083487; PMID:10617197
A/Accession: G84564
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-508 <STO>
A/Cross-references: GB:AB002093; NID:g4218010; PIDN:AAD12218.1; GSPDB:GN00139

C;Genetics:
A;Gene: At2g18480
A;Map position: 2
C;Superfamily: glucose transport protein

Query Match 12.1%; Score 493.5; DB 2; Length 508;
Best Local Similarity 22.7%; Pred. No. 6.8e-27;
Matches 170; Conservative 99; Mismatches 174; Indels 305; Gaps 19;

QY 63 AIVASIGNLQGMWNATIAAALVIKKEFQLONEPTVE--GLIVMSLIGATIVTTFSG 119
Db 26 AIVASIIISIFGYDTGVMGSAQIFIRDLKI-NDTQIEVLAGILNLCALVGSLT---AG 80
QY 120 PLSDSIGRRPMLILSSILYFSGILMLWSPNVYLLARFVDFGIGLAATVLPYISET 179
Db 81 KTSVDVIGRRYTIASAVIFLVGSLMGYPNPVLMVGRCTAGVGVFALMIAPVSAEI 140
QY 180 APSEIRGLNLTLPQFSGSGGMFL---SYCMVFGMSLSPDPDRIMLGVLAIPLSF-FG 234
Db 141 SSASHRGFLTSLPELCSIGILGVSNYC--FG-KJTLKLGWRMLMGIAAFPSLILAFG 197
QY 235 LTIFFYLPESPRMLVSKGRMAEAKVLOKLRKDDVSGELSLLEGLVGGDTSEIYTIIG 294
Db 198 IT--RMPESPRMLVMQGRLEAKIM----- 221
QY 295 PATEAADLVTDGDKQITLYGPEEGQSWIARPSKG?IMLGSVLSLASRHGSMVQSVL 354
Db 222 ----- 221
QY 355 MDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDEEYAS 414
Db 222 -----VLVSNTEEEAEER----- 234
QY 415 DGAGDYEDNLHSPLLSRQATGAEGKDIVHGHRSALSMRQTLLEGGDVSSSTDIGG 474
Db 235 -----FRDIL-----TAAB-----VDVTEIKEVGG 254
QY 475 GMQLAMKWKSEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDVFEGSEFVHAAA 534
Db 255 G-----VKKNHKK----- 263
QY 535 LVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLF---EPGVRALLVGVGIQLQ 591
Db 264 -----SVRELVIKPRPAVRLLILIAAVGIHFFE 291
QY 592 QFAGINGVLYTTPQILEQAGVA---VILSKFGLSSASASILISSLTLLMPCIGFAML 647
Db 292 HATGIEAVVLYSPRIFFKAGVSKDKLLLATVGVGLTKAFPII-----IATF 338
QY 648 LMDLSGRFLLLTGIPILLASVILVNSLI--DLGTLAHAL-LSTVSVIVVFCCEVMGF 704
Db 339 LLDKVGRRKLLLTSTGMMVFALTSLAVSLTMVQRFGLAWALSLISTVYAFVAFSIGL 398
QY 705 GPIPNILCAEIPFTRVRGLCIAICFTWIGDIIIVTSLPVMNAIGLAVFSIYAVVCL 764
Db 399 GPITWVYSSEIFPLRLRAQASIGVAVNRIMNATVSMGFLSMTKAITTGCVFVFAGIAV 458
QY 765 ISFVFEVLKVPETKGMPLLEVITEFFAVG 792
Db 459 AAWMFFFFMLPETKGLPLEMEKLFGGG 486

RESULT 13
A85433
sugar transporter like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C;Accession: A85433
R;anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: A85433

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: GB:NC_001268; NID:g7270615; PIDN:CAB80333.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g36670
A;Map position: 4
C;Superfamily: glucose transport protein

Query Match 12.1%; Score 493; DB 2; Length 493;
Best Local Similarity 23.2%; Pred. No. 7.1e-27;
Matches 177; Conservative 102; Mismatches 178; Indels 306; Gaps 23;

QY 48 SDDILEDKMSG---AVLVAIVASIGNLQGMWNATIAAALVIKKEFQLONEPTVE--- 100
Db 2 ADQISGEKPAGVNRPALQCAIVASIVSIFGYDTGVMGSAVFIEEDLK-TNDVQIEVL 60
QY 101 GLIVMSLIGATIVTTFSGPLSDSIGRRPMLILSSILYFSGILMLWSPNVYLLARFV 160
Db 61 GILNLCALVGSLL-----AGRTSDIIGRRYTIYLAISILFMLGSLIMGWGNYPVLLSGRCT 116
QY 161 DFGIGLAATVLPYISETIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSP-DWR 219
Db 117 AGLGVFALMVAPVYSAEIATASHRGLASLPHLCISIGILGYIVNYPFSKLPMTIGWR 176
QY 220 IMLGVLAIPSLFF-FGLTIFFLPESPRMLVSKGRMAEAKVLOKLRKDDVSGELSLLE 278
Db 177 IMLGIAVPSLVLAFG--ILKMPESPRMLMQGRLEKEKIL----- 216
QY 279 GLEVGGDTSEIYTIIGPATEAADLVTDGDKQITLYGPEEGQSWIARPSKPIMLGSVL 338
Db 217 -----ELVSN-----SPEEAE----- 227
QY 339 SLASRHGSMVQSVPLMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSVTDQAKNE 398
Db 228 -----LR----- 229
QY 399 QWDEENLHRDEEYASDGAGDYEDNLHSPLLSRQATGAEGK---DIVHGHRSALSMR 455
Db 230 -----FQD-----IKAAAGIDPKCVDVV----- 248
QY 456 RQTLLEGGDVSSSTDIGCGQLAMKWKSEKEGNGRKEGFKRVYLHQEGVPGSRGSIIV 515
Db 249 -----KMEGKK-----THGEV----- 260
QY 516 SLPGGDVFEGSEFVHAAALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLF-- 573
Db 261 -----WKELILR 267
QY 574 -EPGVRALLVGVGIQLQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISS 632
Db 268 PTPAVRRVLLTALGIHFQHASGIEAVLYGPRIFFKAGITT-KDKLPLVTIIGVIM--- 323
QY 633 LTTLLMPCIGFAMLLMDLSGRFLLLTGIPILLASVIL-----VVSNLIDLGLTAHAL 687
Db 324 KTTFIPT-----ATLLDKVGRKLLTSVGMVIALTMLGFLGTMAQNA--GGKLAWAL 376
QY 688 -LSTVSVIVVFCCEVMGFPIPNILCAEIPFTRVRGLCIAICFTWIGDIIIVTSLPVM 746
Db 377 VLSIVAAYSFAVAFSIGLGPITWVYSSEVFPDLKRAQASLGVAVNRVNMATVSMGFLSL 436
QY 747 LNAIGLAVFSIYAVVCLISFVFEVLKVPETKGMPLLEVITEFF 789
Db 437 TSAITTGAFMFAGVAAVAVAMNFFFLPETKGSLEIEALF 479

RESULT 14
G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84864
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

Db	395	WAIALCITM-VCAVVASFISIGLGPITWVYSSEVFPLRLRAOCTSMGV-AVNRVSGVIST	452
QY	741	YSLPVMUNAIGLAGVFSIYAVVCLISFVFVFLKVPETKGMPLLEVITEPF	789
Db	453	FPLP-LSHKITTGGAFFLPGGIAIITAMFFFLTFLPETRGRTLENMHELF	500

Search completed: January 5, 2004, 18:58:44
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 15:50:52 ; Search time 24 Seconds

(without alignments)
1567.557 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSWLAVQTPPTPDLDR.....PLEVITTEFAVGAKQAAMAKA 800

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	527	12.9	461	1	CSBC_BACSU
2	486.5	11.9	482	1	YFIG_BACSU
3	483	11.9	457	1	XYLT_LACBR
4	473.5	11.6	464	1	ARAE_BACSU
5	462	11.3	612	1	ITR2_YEAST
6	451	11.1	464	1	GALP_ECOLI
7	434.5	10.7	472	1	ARAE_KLEBOX
8	434	10.7	491	1	XYLE_ECOLI
9	432	10.6	557	1	ITR2_SCHPO
10	431	10.6	547	1	GTR1_LEIDO
11	429.5	10.5	472	1	ARAE_ECOLI
12	421.5	10.3	584	1	ITR1_YEAST
13	421	10.3	629	1	MYCT_HUMAN
14	418.5	10.3	468	1	GLCP_SYNY3
15	404.5	9.9	575	1	ITR1_SCHPO
16	384	9.4	533	1	GTR2_CHICK
17	382	9.4	534	1	HUP3_CHLKE
18	364.5	8.9	570	1	HXT1_YEAST
19	354	8.7	492	1	GTR1_BOVIN
20	352.5	8.7	541	1	HXT2_YEAST
21	352	8.6	492	1	GTR1_MOUSE
22	352	8.6	522	1	STP1_RICCO
23	350.5	8.6	478	1	GTR8_RAT
24	350	8.6	477	1	GTR8_MOUSE
25	350	8.6	492	1	GTR1_RAT
26	349.5	8.6	592	1	HXT5_YEAST
27	348.5	8.6	522	1	GTR2_RAT
28	348.5	8.6	523	1	GTR2_MOUSE
29	348	8.5	492	1	GTR1_HUMAN
30	346.5	8.5	534	1	HUP1_CHLKE
31	346.5	8.5	574	1	GAL2_YEAST
32	346	8.5	524	1	GTR2_HUMAN
33	345.5	8.5	546	1	GHT5_SCHPO

34	345	8.5	490	1	GTR1_CHICK	P46896 gallus gall
35	344	8.4	509	1	GTR4_HUMAN	P14672 homo sapien
36	343	8.4	477	1	GTR8_HUMAN	Q9AY64 homo sapien
37	343	8.4	566	1	KHT2_KLULA	P53387 kluyveromyc
38	343	8.4	567	1	HXT3_YEAST	P32466 saccharomyc
39	342	8.4	492	1	GTR1_RABIT	P13355 oryctolagus
40	342	8.4	496	1	GTR3_CHICK	P28568 gallus gall
41	342	8.4	570	1	HXT6_YEAST	P39003 saccharomyc
42	342	8.4	570	1	HXT7_YEAST	P39004 saccharomyc
43	341	8.4	576	1	HXT4_YEAST	P32467 saccharomyc
44	339	8.3	507	1	GTR6_HUMAN	Q9UGQ3 homo sapien
45	339	8.3	522	1	STP1_ARATH	P23586 arabidopsis

ALIGNMENTS

RESULT 1
CSBC_BACSU STANDARD; PRT; 461 AA.
ID CSBC_BACSU
AC P46333; Q32289; 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable metabolite transport protein csbc.
GN CSBC OR SS92BR.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=96093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
genome between the gnt and tol operons."
RL DNA Res. 2:61-69(1995).
RN [2]
RP REVISIONS.
RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384337;
RA Kunet F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hujlo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tobato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;


```

RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99303315; PubMed=10376822;
RA Akbar S., lee S.Y., Boylan S.A., Price C.W.;
RT "Two genes from Bacillus subtilis under the sole control of the
RT general stress transcription factor sigmaB.";
RL Microbiology, 145:1069-1078(1999).
CC -1- FUNCTION: COULD SERVE EITHER A NUTRITIONAL OR AN OSMOTIC
CC PROTECTION FUNCTION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR	EMBL; AB005554; BAA21604.1; -.
DR	EMBL; Z59124; CAB16017.1; -.
DR	PIR; D70073; D70073.
DR	Subtilist; BG11360; csbC.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR005828; Sub_transporter.
DR	InterPro; IPR005829; Sug_transporter.
DR	InterPro; IPR003663; Sugar_transpt.
DR	Pfam; PF00083; sugar tr; 1.
DR	PRINTS; PR00171; SUGRTRANSPORT.
DR	TIGRfam; TIGR00879; SP; 1.
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW	Transpott; Transmembrane; Complete proteome.
FT	DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 15 35 1 (POTENTIAL).
FT	DOMAIN 36 38 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 39 59 2 (POTENTIAL).
FT	DOMAIN 60 76 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 77 97 3 (POTENTIAL).
FT	DOMAIN 98 104 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 105 125 4 (POTENTIAL).
FT	DOMAIN 126 139 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 140 160 5 (POTENTIAL).
FT	DOMAIN 161 163 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 164 184 6 (POTENTIAL).
FT	DOMAIN 185 241 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 242 262 7 (POTENTIAL).
FT	DOMAIN 263 280 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 281 301 8 (POTENTIAL).
FT	DOMAIN 302 308 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 309 329 9 (POTENTIAL).
FT	DOMAIN 330 341 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 342 362 10 (POTENTIAL).
FT	DOMAIN 363 378 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 379 399 11 (POTENTIAL).
FT	DOMAIN 400 402 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 403 423 12 (POTENTIAL).
FT	DOMAIN 424 461 CYTOPLASMIC (POTENTIAL).
FT	CONFLICT 400 401 RP -> SA (IN REF. 3).
SEQUENCE	461 AA; 50235 MW; B6587B5DC7272E71 CRC64;

Query Match	12.9%;	Score 527;	DB 1;	length 461;
Best Local Similarity	21.5%;	Pred. No. 6.4e-27;		
Matches 156;	Conservative 101;	Mismatches 170;	Indels 300;	Gaps 12;
QY	61	LVAIVASIGNLLQGWMDNATTAALVLYIKKEFQLQNEPFEVSGLIVSMSLIGATIVTFESGP	120	
Db	9	MIYFEGALGGLLYGYDDIGVISCALLFINNDIPLTT-LTEGLIVSMLLIGAFGSALSGT	66	

QY		121	LSDSIGRRPMLILSSILYFFSGGLIMLMSPNVYLRLARFVDGFGIGLAATLVLPYI	SEIA	180
Db		67	CSDRWGRKKVFVLIIIFIGALCAFSQTIGMLIASRVIILGLAVGGSTALVPVYLSEMA		126
QY		181	PSEIRGLINTLTPQFSSGGMFLSYCMWFGMSLSPPDWRIMLGVALPSLFFEGLTIFYL		240
Db		127	PTKIRGTLLGTNNLMIVTGIILAY--IVNYLFTPEAWRMWGAAVPAVALL-IGIAFM		183
QY		241	PESPRWLVSCKGRMAEAKVKLOKLGRKDVSSELSSLLEGLEVGGDTSIEEYIIGPATEA		300
Db		184	PESPRWLVKRGSEBEARRIMNITHPKDIEMELAEWKQ-----		222
QY		301	DDLVTDDKEQITLYGPPEGQSNIARPSKGPIMLGSVLSLASRHGSMVNQSVPLMDPIVT		360
Db		223	-----EAKKKETTL-----		231
QY		361	LFSGVHENMPQAGGSMRSTLFPNFSGMFSVTQHAKNQOWDEENLHRDDEYASDGAGD		420
Db		232	-----		231
QY		421	YEDNLHSPLLSRQATGAEGKDIVHHGRGSALSMRROTLLGEGGDVSSSTDIGGWQLAW		480
Db		232	-----GV-----LKA		236
QY		481	KMSEKEGENGRKEGGFKRVYLHQEGVPCSRGSI VSLPGGDVFEGSEFYHAALLVSQSA		540
Db		237	KW-----		238
QY		541	LFESGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQLIQFAGINGVL		600
Db		239	-----IRPMLLIGVGLAIPOQAVGINTVI		262
QY		601	YYTFQILLEQAGVAVILSKFGLSSASASILSSLTLLMLPCI GFAMLMDUGRRFELL-		659
Db		263	YVAP-----TI FTKAGLGTSASALGTMGIGILNVIMCI-TAMILDRVGRKKLIW		312
QY		660	GTFIPILLASLVILVSNLIDGLTAH-ALLSTVSIVVYFCCFVMGFGPIPINILCAEIEPT		718
Db		313	GSVGTLLSLALSGV-LTLTGLSASTAMTWVFLGVYIVFYQATWGPVAVMLPELFP		370
QY		719	RVRGLCIAICAFTFWIGDITVYSBPMLNAIGLAGVESIYAVVCLISFVFVLKVPETK		778
Db		371	KARGAATGFTTLVLSAANLIVSLVFPMLLRPMGIAMVFMVFSVICLSFFFAFYMVPETK		430
QY		779	GMPLEVI	785	
Db		431	GKSLEEI	437	

RESULT 2			
	YFIG_BACSU	STANDARD;	PRT; 482 AA.
ID	YFIG_BACSU		
AC	P54723;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Hypothetical metabolite transport protein yFIG.		
GN	yFIG.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168;		
RX	MEDLINE=96262713; PubMed=8704981;		
RA	Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;		
RT	"Determination of a 12 kb nucleotide sequence around the 76 degrees		
RL	region of the Bacillus subtilis chromosome.";		
RN	Microbiology 142:1417-1421(1996).		
RP	[2]		
RC	SEQUENCE FROM N.A.		
RX	STRAIN=168;		
RA	MEDLINE=98044033; PubMed=9384377;		
RT			
RL			
RN			
RP			
RC			
RX			

RA	Kunst F, Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA	Borries R., Boursier L., Brans A., Braun M., Brignelli S.C., Bron S.,
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA	Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA	Fritz C., Fujita P., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA	Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA	Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA	Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA	Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA	Parro V., Pohl T.M., Portecelle D., Porwollik S., Prescott A.M.,
RA	Prescan E., Pulic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA	Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA	Viart A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA	Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT	"The complete genome sequence of the Gram-positive bacterium <i>Bacillus</i>
RT	<i>subtilis</i> ."
RL	Nature 390:249-256 (1997).
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; D50543; BAA09111.1; -
DR	EMBL; Z99108; CAB12655.1; -
DR	PIR; B69803; B69803.
DR	Subtilist; BG1854; yfig.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR005828; Sub_transporter.
DR	InterPro; IPR005829; Sug_transporter.
DR	InterPro; IPR003663; Sugar_transpt.
DR	Pfam; PF00083; sugar tr; 1.
DR	PRINTS; PR00171; SUGRTRANSPORT.
DR	TIGRFAMS; TIGR00879; SP; 1.
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW	Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT	DOMAIN 1 29
FT	TRANSMEM 30 50 1 (POTENTIAL).
FT	DOMAIN 51 59 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 60 80 2 (POTENTIAL).
FT	DOMAIN 81 92 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 93 113 3 (POTENTIAL).
FT	DOMAIN 114 120 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 121 141 4 (POTENTIAL).
FT	DOMAIN 142 155 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 156 176 5 (POTENTIAL).
FT	DOMAIN 177 184 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 185 205 6 (POTENTIAL).
FT	DOMAIN 206 263 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 264 284 7 (POTENTIAL).
FT	DOMAIN 285 301 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 302 322 8 (POTENTIAL).
FT	DOMAIN 323 331 CYTOPLASMIC (POTENTIAL).

Query Match	Best Local Similarity	Score	DB 1;	Length	482;
Matches	158;	Conservative	109;	Mismatches	195;
				Indels	303;
				Gaps	15
FT TRANSMEM	332	352	9 (POTENTIAL).		
FT TRANSMEM	353	373	10 (POTENTIAL).		
FT DOMAIN	374	400	CYTOPLASMIC (POTENTIAL).		
FT TRANSMEM	401	421	11 (POTENTIAL).		
FT DOMAIN	422	423	EXTRACELLULAR (POTENTIAL).		
FT TRANSMEM	424	444	12 (POTENTIAL).		
FT DOMAIN	445	482	CYTOPLASMIC (POTENTIAL).		
SEQ SEQUENCE	482 AA;	52756 MW;	90851C4F4C48EE01 CRC64;		
Query Match	11.9%;	Score 486.5;	DB 1;	Length 482;	
Best Local Similarity	20.7%;	Pred. No. 2.9e-24;			
Matches	158;	Conservative 109;	Mismatches 195;	Indels 303;	Gaps 15
QY	39	SCSSQEPVTSDDILEDKMGSAV-LVAIVASIGNLLQGWDNATIAAVALYIKKEFQLONEP	97		
Db	2	STKKEAVIGKESLAHK-GILRTITLVSTFGGLPGYDTGVINGALPFWATAGQLNLTP	59		
QY	98	TVEGLIVMSLIGATIVTTSPSPDSIGRRPMLSSILYFESGLIMLSPNVYVLLA	157		
Db	60	VTEGLVASSLLGAFAFGAMFGRSLDRHGRKRTILYALLFLIATLCTFSPNASVMIAF	119		
QY	158	RFVDFGIGLAVTLVPLYISEIAPSEIRGLNLTLPQFSGSGGMFLSYCM--VFGMSLSPS	215		
Db	120	RFLGLAVGCASVTVPFLAETSPARRGRIVTQNELMIVIGQLLATFNALIGSTMGES	179		
QY	216	PD-WRIMLGVAIPS-LFFEGLTIFLLPESPRVLVSKGRMAEAKKVLQKRGKDDVSGEL	273		
Db	180	ANVMRYMLVIATLPAAVLWFGMLT--VPESPRWLAAKGRMGDALRYLROIR-----	228		
QY	274	SLLLEGLEVGSDTSIEEYIIGPATEADDLVTDGKEQITLYGPEEGSWIARPSKGPIM	333		
Db	229	-----EDSQAQQ-----	235		
QY	334	LGSVLSLASRHGSMVNSQVPLMDPIVTLFGSVHENMPQAGSMRSTLPPNFGSMFSTVDQ	393		
Db	236	-----	235		
QY	394	HAKNEQWDEENLHRDDEEYASDAGDYEDNLHSPLSRQATGAEGKDIVHGHGHSALS	453		
Db	236	-----EIKEIKH-----	242		
QY	454	MRRQTLLEGGDVSTDIGGQWLAKWKEKEGENGRKEGFKRYVYHQEGVPGSRGS	513		
Db	243	-----	242		
QY	514	IVSLPGGDVFESEFVHAALVQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLF	573		
Db	243	-----AIEGT-----AKKAGFHDQ	257		
QY	574	EPGVRALLVGVGIQTLQGFAGINGVLVYTPQILEQAGVAVILSKFGLSSASASILSSL	633		
Db	258	EPWIRKILFIGIGIAIVQQTIGVNSIMYYGTEILREAGF-----QTEALLIGNIA	307		
QY	634	TTLLMLPCIGFAMLLMDLSGR-FLLLGTIPILIASLVILVNSNLIDGTLAALLSTVS	692		
Db	308	NGVISVIAVIFGIWLLGKVRRRPMLIQIGTMTALLIGLISYLE-CTPALPYV-VLS	365		
QY	693	VIVYFCCFVM-GFGPIPNILCAETPTVRGLCIAICAFTFWIGDIIVYVSLPVMNAIG	751		
Db	366	LTLFLAFOQTALSTVTWMLSELFPMHVRLGIMGISTFCLMTANFLIGTTPILLNHIG	425		
QY	752	LAGVFSIYAVVCLISFVFVFLKVPETKGMPLLEVITEFFAVGAKQA	796		
Db	426	MSATFFIFVAMNITALLFVKKYVPETKGRSLSEQLEHSFRQYGRRA	470		
RESULT 3					
XYLT_LACBR	STANDARD;	PRT;	457 AA.		
AC	052733;				
DT	15-DEC-1998	(Rel. 37, Created)			
DT	15-DEC-1998	(Rel. 37, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			

```

DE D-xylose-proton symporter (D-xylose transporter).
GN XYL1.
OS Lactobacillus brevis.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1580;
RN [1]
RP SEQUENCE FROM N.A.
RA Chailion S., Bor Y.-C., Batt C.A., Postma P., Pouwels P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: UPTAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

RP SEQUENCE FROM N.A.
RA Challou S., Bor Y.-C., Batt C.A., Postma P., Pouwels P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBD databases.
CC -!- FUNCTION: UPTAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF045552; AAC95127.1; -.
CC DR InterPro; IPR007114; MFS.
CC DR InterPro; IPR005828; Sub_transporter.
CC DR InterPro; IPR005829; Sug_transporter.
CC DR InterPro; IPR003663; Sugar_transpt.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGRTRANSPORT.
CC TIGRfams; TIGR00879; SP; 1.
CC DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
CC KW Transport; Sugar transport; Transmembrane; Symport.
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 14 34 1 (POTENTIAL).
FT TRANSMEM 35 45 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 46 66 2 (POTENTIAL).
FT DOMAIN 67 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 101 3 (POTENTIAL).
FT DOMAIN 102 103 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 104 124 4 (POTENTIAL).
FT DOMAIN 125 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 151 5 (POTENTIAL).
FT DOMAIN 152 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 184 6 (POTENTIAL).
FT DOMAIN 185 243 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 244 264 7 (POTENTIAL).
FT DOMAIN 265 280 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 301 301 8 (POTENTIAL).
FT DOMAIN 302 309 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 310 330 9 (POTENTIAL).
FT DOMAIN 331 337 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 338 358 10 (POTENTIAL).
FT TRANSMEM 359 379 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 380 400 11 (POTENTIAL).
FT DOMAIN 401 401 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 402 422 12 (POTENTIAL).
FT DOMAIN 423 457 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 457 AA; 49199 MW; E097EB2B67A92F67 CRC64;

```

Query Match	11.9%	Score 483;	DB 1;	Length 457;
Best Local Similarity	21.3%;	Pred. No. 4.5e-24;		
Matches 155; Conservative	95;	Mismatches 177;	Indels 300;	Gaps 14;

```
Qy      62 VAIVASIGNLLQGWDNATIAAAVLVIKKEPOLONEPTVEGLIVSMSLIGATTVTFSGPL 121
        ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      9 VYFEGALGGLFGYDTGVISGALFIQKNMNGS--WQQGWVSVAVLLGAILGAALIGPS 66
Qy      122 SDSIGRPMILSSILVFSSGLIMLMSPNVVLLARFVDGFIGLAVTLVPLISEIAP 181
        |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      67 SDFGRKKULLLSAIIFFVGALGSASFSPPEWTLIISRITLGAVGAASALIPTYLAELAP 126
```

```

QY      182 SEIRGLINTLPQFSGSGGMFLSYCMVEGMSLSPPDWRIMLGVAIPS--LFFFGLTIFY 239
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      127 SDKRGTVSSLFQLMVMGTGILLAYITNYSFS-GFYTGWRMMLGFAIPALLFLGGL---I 182

QY      240 LPESPRLVSKGRMAEAKKYLQKLRGKDDVSGELSLLLGLEGVGDTSEETIIGPATEA 299
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      183 LPESPRLVSKGHLDEARHVLDTMNKHDOVA----- 213

```

QY 300 ADDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSLASRHGSMVNQSVPLMDPIV 359

Db 214 -----VNKEI----- 218

27 300 14103 VIENTNE ZROGOWNS 141 FNF OSNF SV 1 DQHNNEQWDEENINWDEE INDDGAG 713

426 ПУЕДНУ ИЕДИ І СРОДТЕСА ЕКІДІТІУШУ СІУБСА І СМІДІАТІ І СЕССІДІУ СЕДІТІ ССІСІАІ І 426

```

Db      224 -----SAXIVSGW----- 232
          | : : |||

```

QY 480 WKWSEKEGNGRIKEGFKRVYLHQEGVPGSRGSI VSLPGGDVFEGSEFVHAALVSQS 539

Db 233 -----S 233

540 ALFSKGLAEPKMSDAAMVHPSEVAKGSRWKDLFEPGVRALLVGVIQLQDFAGINGV 599

Country	Year	Population (millions)	Urban population (millions)	Urban population (%)	Population density (per sq km)	Urban population density (per sq km)
Algeria	1980	10.5	4.5	42.9	105	105
Algeria	1985	11.5	5.5	47.8	115	115
Algeria	1990	12.5	6.5	51.6	125	125
Algeria	1995	13.5	7.5	55.5	135	135
Algeria	2000	14.5	8.5	58.6	145	145
Algeria	2005	15.5	9.5	61.3	155	155
Algeria	2010	16.5	10.5	63.6	165	165
Algeria	2015	17.5	11.5	65.7	175	175
Algeria	2020	18.5	12.5	67.6	185	185
Algeria	2025	19.5	13.5	69.2	195	195
Algeria	2030	20.5	14.5	70.7	205	205
Algeria	2035	21.5	15.5	72.1	215	215
Algeria	2040	22.5	16.5	73.3	225	225
Algeria	2045	23.5	17.5	74.5	235	235
Algeria	2050	24.5	18.5	75.5	245	245
Algeria	2055	25.5	19.5	76.5	255	255
Algeria	2060	26.5	20.5	77.4	265	265
Algeria	2065	27.5	21.5	78.2	275	275
Algeria	2070	28.5	22.5	78.9	285	285
Algeria	2075	29.5	23.5	79.7	295	295
Algeria	2080	30.5	24.5	80.3	305	305
Algeria	2085	31.5	25.5	81.0	315	315
Algeria	2090	32.5	26.5	81.6	325	325
Algeria	2095	33.5	27.5	82.1	335	335
Algeria	2100	34.5	28.5	82.6	345	345
Algeria	2105	35.5	29.5	83.1	355	355
Algeria	2110	36.5	30.5	83.6	365	365
Algeria	2115	37.5	31.5	84.0	375	375
Algeria	2120	38.5	32.5	84.4	385	385
Algeria	2125	39.5	33.5	84.8	395	395
Algeria	2130	40.5	34.5	85.2	405	405
Algeria	2135	41.5	35.5	85.5	415	415
Algeria	2140	42.5	36.5	85.9	425	425
Algeria	2145	43.5	37.5	86.2	435	435
Algeria	2150	44.5	38.5	86.5	445	445
Algeria	2155	45.5	39.5	86.8	455	455
Algeria	2160	46.5	40.5	87.1	465	465
Algeria	2165	47.5	41.5	87.4	475	475
Algeria	2170	48.5	42.5	87.7	485	485
Algeria	2175	49.5	43.5	88.0	495	495
Algeria	2180	50.5	44.5	88.3	505	505
Algeria	2185	51.5	45.5	88.5	515	515
Algeria	2190	52.5	46.5	88.8	525	525
Algeria	2195	53.5	47.5	89.0	535	535
Algeria	2200	54.5	48.5	89.2	545	545
Algeria	2205	55.5	49.5	89.4	555	555
Algeria	2210	56.5	50.5	89.6	565	565
Algeria	2215	57.5	51.5	89.7	575	575
Algeria	2220	58.5	52.5	90.0	585	585
Algeria	2225	59.5	53.5	90.1	595	595
Algeria	2230	60.5	54.5	90.2	605	605
Algeria	2235	61.5	55.5	90.4	615	615
Algeria	2240	62.5	56.5	90.5	625	625
Algeria	2245	63.5	57.5	90.6	635	635
Algeria						

Db 263 IYYAPTETDWC-----EGVSAAILAHIGIGTENVIV---TAIAVAIMDKRIDPKKTYN 312

QY 659 LGTIPILIASLVILVVSNLIDGLIAHALLSTVSIVVFCFCFVMGFGPIPNILCAEIPPT 718

Db 313 IGAVGMGI-SLFVMSIGMKFSGGSQTAAISVIALTVYIAFFSATWGPVMMVMI GEVPL 371

719 RVRGLCIAICAFTEWIGDIIVTYSLPVMNAIGLACVFSIYAVVCLISFVFVFLKVPETK 778

DB 3 / 2 NIKGUGNSFASV INWIANMIVSLIFPSLEDFGIGSUF IGIGLCCFASIMF VQKNVEIK 431

[illegible]

ARAE_BACSU

AC P96710;
 26 FEB 2003 (F03 41 C200404)

DT	28-FEB-2003	(Rel. 41, Last sequence update)
DT	28-FEB-2003	(Rel. 41, Last annotation update)

GN ARAE.
DE PROCON SYMPOTECI (HABITABO CUMPROTECI).
DE HABITABO CUMPROTECI.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI TaxID=1432.

RN	[1]
RP	SEQUENCE FROM N.A.

RX MEDLINE=98044033; PubMed=9384377;

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borrás R., Boursier I., Brans A., Braun M., Brinnell S.C., Bron S.

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 Broillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guissem J., Guy B.T., Hada K., Hajech J., Harwood C.B., Henaut A.

KA Hildert H., Holsappel S., Hosono S., Huilo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., RA

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Melado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portollik S., Prescott A.M.,
RA Pressecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambolt R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.";
RL Nature 390:249-256 (1997).
RN [2]
RP SEQUENCE OF 1-223 FROM N.A.
RC STRAIN=168;
RX MEDLINE=97197523; PubMed=9045819;
RA Sa-Nogueira S., Mota L.J.;
RT "Negative regulation of L-arabinose metabolism in *Bacillus subtilis*:
RT characterization of the *araC* (araC) gene.";
RL J. Bacteriol. 179:1598-1608 (1997).
RN [3]
RP SEQUENCE OF 223-464 FROM N.A., AND FUNCTION.
RC STRAIN=168;
RX MEDLINE=98062200; PubMed=9401028;
RA Sa-Nogueira I., Ramos S.S.;
RT "Cloning, functional analysis, and transcriptional regulation of the
RT *Bacillus subtilis* *araE* gene involved in L-arabinose utilization.";
RL J. Bacteriol. 179:7705-7711 (1997).
RN [4]
RP TRANSCRIPTIONAL REGULATION.
RX MEDLINE=99348379; PubMed=10417639;
RA Mota L.J., Tavares P., Sa-Nogueira I.M.G.;
RT "Mode of action of *AraC*, the key regulator of L-arabinose metabolism
RT in *Bacillus subtilis*.";
RL Mol. Microbiol. 33:476-489 (1999).
CC -!- FUNCTION: Uptake of arabinose across the boundary membrane with
CC the concomitant export of a proton (symport system).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- INDUCTION: Transcription is repressed by the binding of *araC* to
CC the promoter. L-arabinose acts as an inducer by inhibiting the
CC binding of *araC* to the DNA, thus allowing expression of the gene.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z99121; CAB15401.1; -;
DR EMBL; X98354; CA66998.1; -;
DR EMBL; Y12105; CA72812.1; -;
DR PIR; F69587; F69587.
DR Subtilist; BG11907; araE.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Symport; Complete proteome.

FT TRANSMEM 21 43 POTENTIAL.
FT TRANSMEM 63 85 POTENTIAL.
FT TRANSMEM 92 111 POTENTIAL.
FT TRANSMEM 116 138 POTENTIAL.
FT TRANSMEM 150 172 POTENTIAL.
FT TRANSMEM 185 207 POTENTIAL.
FT TRANSMEM 266 288 POTENTIAL.
FT TRANSMEM 303 325 POTENTIAL.
FT TRANSMEM 332 354 POTENTIAL.
FT TRANSMEM 364 386 POTENTIAL.
FT TRANSMEM 398 420 POTENTIAL.
FT TRANSMEM 424 446 POTENTIAL.
SQ SEQUENCE 464 AA; 50411 MW; 13B417061CB61DA2 CRC64;

Query Match 11.6%; Score 473.5; DB 1; Length 464;
Best Local Similarity 20.0%; Pred. No. 1.9e-23;
Matches 153; Conservative 102; Mismatches 191; Indels 319; Gaps 13;

QY 37 PASCSSQEPVTSDDILEDKMSGAVLVAIVASIGNLLQGDWNTAIAAALYIKKEFQONE 96
Db 5 PTQLEPNVPVTR---SHSMGFVILISCAAGLGILGYDTAVISGAIGFLKDYSL--S 58

QY 97 PVEGLIVMSLIGATVITTFSGPLSDSIGRRPMLLSSILYFSGLIMLSPNYVLL 156
Db 59 PMEGLVSSIMIGVGVGISGFLSDRFGRRKIMTAALLFAISAVISALSDVSTLLI 118

QY 157 ARFVDGFIGLAVTVLPYLISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSP 216
Db 119 ARIIGGLIGMGSSLSVYITEAPPAIRGSLSLYQLFTILGISATYFINLAQRSGTY 178

QY 217 D-----WRIMLVLAIPSLFFPGLTIFYLPSPRMLVSKGRMAEAKVLQKRGKDV 270
Db 179 EWGVHTGWRMLAYGVPSVIFP-LVLLVPESPRLAKAGKTNEALKILTRINGETVAK 237

QY 271 GELSLLEGLVEVGDTISIEYIIGPATEAADLVTDGKEQITLYGPEEGSWIARPSKG 330
Db 238 EELKNIENSLKI-----EQM----- 252

QY 331 PIMLGSVLSASRHGSMVNSVPLMDPIVTLFGSVHNMPPQAGSMRSTLPNFGSMFSV 390
Db 253 ----- 252

QY 391 TDQAKNEQWDEENLRHDEEYASDAGGDYEDNLHSPLLSRQATGAEGKDIYHGHRS 450
Db 253 ----- 252

QY 451 ALSMRQTLLEGGDVSTDIGGQWLAKWSEKGENGRKEGFKRVYLHQEGVPSR 510
Db 253 ----- 252

QY 511 RGSIVSLPGGDVFEGSEFVHAALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWK 570
Db 253 -GSL-----S 256

QY 571 DLFEQVRRLAVGVIQILQFAGINGVLVYTPQILEQAGVAVILSKFGLSASASILI 630
Db 257 QLEKPKLRKALVIGILALFNQVIGMNAITYGPELFKMG-----FGQNAQFVYTCI 309

QY 631 SSLTLLMLPCIGFAMLLMDLSGRFL-LGTIPILASLVILVNSLIDLT-----LA 684
Db 310 VGVEVEVFTV---IAVLIDKVRKKLMISGS--AFMAIFMILI-----GTSFYFELT 357

QY 685 HALISTSVIVVYFCCFWGFGPIPNILCAEIPFTRVGLCIAICAFTFWIGDIITYSLP 744
Db 358 SGIMIVLILGVAAFCVSVGPITWIMISEIFPNHARARAAGIATIFLWGANAIGQFVP 417

QY 745 VMINAIGLAVFSIYAVVCLISFVFLKVPETKMPLEVITEPF 789
Db 418 WMIDSFGLAYTFWIFAVINILCFLFVVTICPETKNSHEIEIKLW 462

RESULT 5
ITR2_YEAST

ID	GALP	ECOLI	STANDARD;	PRT;	464	AA.
AC	P37021;					
DT	01-JUN-1994	(Rel. 29, Created)				
DT	01-JUN-1994	(Rel. 29, Last sequence update)				
DT	28-FEB-2003	(Rel. 41, Last annotation update)				
DE	Galactose-proton symporter (Galactose transporter).					
GN	GALP OR B2943 OR C3529.					
OS	Escherichia coli, and					
OS	Escherichia coli O6.					
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
OC	Enterobacteriaceae; Escherichia.					
OX	NCBI_TaxID=562, 217992;					
RN	[1]					
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.					
RA	Roberts P.E.;					
RL	Thesis (1992), University of Cambridge, U.K.					
RN	[2].					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=K12 / MG1655;					
RX	MEDLINE=97426617; PubMed=9278503;					
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,					
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,					
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,					
RA	Mau B., Shao Y.;					
RT	"The complete genome sequence of Escherichia coli K-12.";					
RL	Science 277:1453-1474(1997).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=O6:H1 / CFT073 / ATCC 700928;					
RX	MEDLINE=22388234; PubMed=12471157;					
RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,					
RA	Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,					
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,					
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;					
RT	"Extensive mosaic structure revealed by the complete genome sequence					
RT	of uropathogenic Escherichia coli.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).					
CC	-1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH					
CC	THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).					
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.					
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; U28377; AAA69110.1; -					
DR	EMBL; AE000377; AAC75980.1; -					
DR	EMBL; AE016766; AAN81977.1; ALT_INIT.					
DR	PIR; F65079; F65079.					
DR	EcoGene; EG12148; galp.					
DR	InterPro; IPR007114; MFS.					
DR	InterPro; IPR005828; Sub_transporter.					
DR	InterPro; IPR005829; Sug_transporter.					
DR	InterPro; IPR003663; Sugar_transpt.					
DR	Pfam; PF00083; sugar tr; 1.					
DR	PRINTS; PR00171; SUGRTNSPORT.					
DR	TIGRfams; TIGR00879; SP; 1.					
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.					
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.					
KW	Transport; Sugar transport; Transmembrane; Inner membrane; Symport;					
KW	Complete proteome.					
FT	DOMAIN	1	15			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	16	36			1 (POTENTIAL).
FT	DOMAIN	37	56			PERIPLASMIC (POTENTIAL).
FT	TRANSMEM	57	77			2 (POTENTIAL).
FT	DOMAIN	78	84			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	85	105			3 (POTENTIAL).
FT	DOMAIN	106	112			PERIPLASMIC (POTENTIAL).

Query Match	Best Local Similarity	Score	DB	Length
Matches 160; Conservative	21.6%;	451;	1;	464;
		Pred. No. 5.5e-22;		
		88; Mismatches 183; Indels 310; Gaps 16		
QY	62	VAIVASIGNLLQGWDNATIAAVLYIKKEFQLONEPTVEGLIVSMILGATIVTTFSGPL	121	
Db	18	VCFALALAGLLFGLDIGIAGALPFIADFEQITSH--TQEWVSSMMFGAIVAGVSGWL	75	
QY	122	SDSIGRRPMLILSSILYFESGLIMLSPNVVYLLARFVDGFIGLAVTLVPLYISEIAP	181	
Db	76	SFKLGRKKSMLMIGAILFVAGSLFSAAPNVEVLILSRVLLGLAVGASATAPLYISEIAP	135	
QY	182	SEIRGLINTLPQFSSGGMFLSYCNVFGMSLSPSPDWRIMLGVAIPSLFFFGLTIFYLP	241	
Db	136	EKIRGSMISMYQLMTTIGILGAY--LSDTAFSYTGAMRWMLGVIIIPAILLL-IGVFPLP	192	
QY	242	ESBPRLVSKGRMAEAKKVLQKRGKDVSGELSLLEGLLEVGGDTSEIYYIIGPATEAAD	301	
Db	193	DSPRFWFAKRFRFVDAERVLLRLR-----DTSAE-----	220	
QY	302	DLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLARHGSVMNQSVPLMDPIVTL	361	
Db	221	-----	220	
QY	362	FGSVHENMPQAGGSMRSTLPNFGSMFSVTDQAKNEQWDEENLHRDEEYASDGAAGDY	421	
Db	221	-----AKREL-----	225	
QY	422	EDNLHSPLLSRQATGAEGKDIVHGHGHSALSMRRQTLLEGGDVSSSTDIGGGQLAWK	481	
Db	226	-DEIRESLQVKQS-----GVAL-----	241	
QY	482	WSEKEGENGRKEGGFKRVYLHQEGVPGSRRSIVSLPGGGDVFEESSEFYHAALVSQSAL	541	
Db	242	-----FK-----ENSNF-----	248	
QY	542	FSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIIQLQOPAGINGVLY	601	
Db	249	-----RAVFLGVLLQVMQOPTGMNVIMY	272	
QY	602	YTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSGRFLLIGT	661	
Db	273	YAPKIFELAGYTNTEQMW---GTVIVGLTNVLAIFAIG---LVDRWGRK-----	317	
QY	662	IPILIASLVILVSNLIDIGTLAALLSTVS---VIVYFCFVMGF---GPINILCA	713	
Db	318	-PTLTLGLFLVMAAGMV-LGTMHHIGHSPSAQYFAIAMLLMFIIVGFAMSAGPLIWLCS	375	
QY	714	EIEPFRVRGLCIAICAFTFWIGDIYIYSLPVMUNAIGLAGVFSIYAVVCLISFVFEFLK	773	
Db	376	EIOPIKGRDFGICSTATNWIAMNIVGATFTMLNTLGNANTFWYAAALNVLFILLTLM	435	
QY	774	VPETKGMPLLEVITTEFFAVGAK	794	

Db 436 VPETKVSLEHIERNLKMKGRK 456

RESULT 7

ARAE_KLEOX ID ARAE_KLEOX STANDARD; PRT; 472 AA.

AC P45598;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Arabinose-proton symporter (Arabinose transporter).

GN ARAE.

OS Klebsiella oxytoca.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Klebsiella.

OX NCBI_TaxId=571;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=8017;

RX MEDLINE=95394866; PubMed=7665532;

RA Shatwell K.F., Charalambous B.M., McDonald T.P., Henderson P.J.F.;

RT "Cloning, sequencing, and expression of the araE gene of Klebsiella

RT oxytoca 8017, which encodes arabinose-H⁺ symport activity.";

RL J. Bacteriol. 177:5379-5380(1995).

CC -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH

CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or_send_email_to_license@isb-sib.ch)

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X79598; CAA56110.1; -.

DR PIR; S47089; S47089.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.

DR InterPro; IPR005829; Sug_transporter.

DR InterPro; IPR003663; Sugar_transpt.

DR Pfam; PF00083; sugar_tr; 1.

DR PRINTS; PRO0171; SUGRTNSPORT.

DR TIGRfam; TIGR00879; SP; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.

FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 30 50 1 (POTENTIAL).

FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 64 84 2 (POTENTIAL).

FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 92 112 3 (POTENTIAL).

FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 115 135 4 (POTENTIAL).

FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 155 175 5 (POTENTIAL).

FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 179 199 6 (POTENTIAL).

FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 258 278 7 (POTENTIAL).

FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 298 318 8 (POTENTIAL).

FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 326 346 9 (POTENTIAL).

FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 362 382 10 (POTENTIAL).

FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 405 425 11 (POTENTIAL).

FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 428 446 12 (POTENTIAL).

FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).

SQ SEQUENCE 472 AA; 51732 MW; 410021E1BEE3D96E CRC64;

Query Match 10.7%; Score 434.5; DB 1; Length 472;

Best Local Similarity 21.4%; Pred. No. 6.6e-21;

Matches 158; Conservative 98; Mismatches 182; Indels 299; Gaps 14;

QY 62 VAIVASIGNLQGMNDATIAAVALYIKKEFQIQNEPTVEGLIVSMSLGATIVTFSGPL 121

Db 25 VSIAAVALGLFGLDIGVIAGALPFTIDHFLVSSR--LQEWVSSMMLGALGALFNGWL 82

QY 122 SDSIGRRPMLILSLYFFSGGLIMLSPNVYVLLARFVDGFGIGLAVTLVPLYSEIAP 181

Db 83 SFRIGRKYSIMVGAVIFVAGSVGSATFATSEMLIVARIVLGAVAGIASYADLYSEMAS 142

QY 182 SEIRGLNTLPQSGSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALIPSLFPEGTLFYLP 241

Db 143 ENVRGKMISMYQLMVTYLGIVMAF--LSDTAFSYSGNWRAMLGVALLPAVLLIIVIF-LP 199

QY 242 ESPRWLVSKGRMAEAKVYLQKRGKDDVSGELSLLEGLVGGDTSEIETIIGPATEAAD 301

Db 200 NSPRWLAEKGRHVEAEVLRMLR-----DTS-----EKARD 230

QY 302 DLVTGDKEQITLYGPEEGQSWIARPSKGPIMGSLVSLASRHGSMVNGSVPLMDPIVTL 361

Db 231 EL-----NEIRESLKQ----- 243

QY 362 FGSVHENMPQAGSMRSTLPNFGSMFSVTDQAKNEQWDEENLHRDDEEYASDGAAGDY 421

Db 244 ----- 243

QY 422 EDNLHSPILSRQATGAEGKDIVHGHGSGALSMRQTLLEGGGDVSTDIGGQWLAWK 481

Db 244 -----GGW----- 246

QY 482 WSEKEGENGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDVFEGSEFVHAALVSQSAL 541

Db 247 -----AL 248

QY 542 FSKGLAEPMSDAMVHPSEVAAGSRWKDLPEPVRALLVGVGIQLQOFAGINGVLY 601

Db 249 F-----KVR-----NVRRAVFLGMLLQMQQFTGMNIIMY 279

QY 602 YTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLMDSGR--FLLLG 660

Db 280 YAPRIFKMA-----GFTTEQQMVATLVGLTFMFATFIIVFVYKAKRPAKLG 330

QY 661 TIPILIASLVILVSNLIDLGLTALHALSTVSVIVYFCC--FVMGFGPLPNILCAEIFP 717

Db 331 FSVMAIGTLVGYCLMQFDNGT--ASSGLSWLSVGMTMCIAGYAMSAPVWILCSEIQP 389

QY 718 TRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLGVFSIYAVVCLISFVFLKVPET 777

Db 390 LKCRDFGITCSTTNWVSNMIIATFTLLDAIGAAGTFWLTYTALNVAFIGVTFWLIPET 449

QY 778 KGMPLVITEFPAVGAK 794

Db 450 KNVTLEHIERLRMSGEK 466

RESULT 8

XYLE_ECOLI ID XYLE_ECOLI STANDARD; PRT; 491 AA.

AC P09058;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE D-xylose-proton symporter (D-xylose transporter).

GN XYLE OR B4031 OR Z5629 OR ECSS014.

OS Escherichia coli, and

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88007632; PubMed=2820984;
RA Davis E.O., Henderson P.J.F.;
RT "The cloning and DNA sequence of the gene xylE for xylose-proton
symport in Escherichia coli K12.";
RL J. Biol. Chem. 262:13928-13932(1987).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=87115869; PubMed=3543693;
RX Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA Henderson P.J.F.;
RT "Mammalian and bacterial sugar transport proteins are homologous";
RL Nature 325:641-643(1987).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RT Daniels D.L.;
RL "Analysis of the Escherichia coli genome. IV. DNA sequence of the
region from 89.2 to 92.8 minutes.";
RN Nucleic Acids Res. 21:5408-5417(1993).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RL "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RN Nature 409:529-533(2001).
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
RN [6]
RN SEQUENCE OF 1-192 FROM N.A.
RP STRAIN=K12;
RX MEDLINE=88234001; PubMed=2836810;
RA Francoz E., Dassa E.;
RT "3' end of the maleFG operon in E.coli: localization of the
transcription termination site.";
RL Nucleic Acids Res. 16:4097-4109(1988).
CC -1- FUNCTION: UPTAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH
THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- INDUCTION: By xylose.
CC -1- MISCELLANEOUS: E.COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT
ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM
WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS
INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYP SYSTEM THAT USES A
HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC
SHOCK.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entitles requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02812; AAA79016.1; -.
DR EMBL; U00006; AAC43125.1; -.
DR EMBL; AE000476; AAC77001.1; -.
DR EMBL; AE005636; AAG59230.1; -.
DR EMBL; AP002568; BAB38437.1; -.
DR EMBL; X06663; CAA29863.1; -.
DR PIR; A26430; A26430.
DR PIR; B86096; B86096.
DR PIR; F91255; F91255.
DR EcoGene; EG11076; xyle.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 30 1 (POTENTIAL).
FT DOMAIN 31 55 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 56 76 2 (POTENTIAL).
FT DOMAIN 77 89 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 90 110 3 (POTENTIAL).
FT DOMAIN 111 133 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 134 154 4 (POTENTIAL).
FT DOMAIN 155 165 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 166 186 5 (POTENTIAL).
FT DOMAIN 187 200 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 201 221 6 (POTENTIAL).
FT DOMAIN 222 272 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 273 293 7 (POTENTIAL).
FT DOMAIN 294 312 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 313 333 8 (POTENTIAL).
FT DOMAIN 334 343 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 344 364 9 (POTENTIAL).
FT DOMAIN 365 369 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 370 390 10 (POTENTIAL).
FT DOMAIN 391 407 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 408 428 11 (POTENTIAL).
FT DOMAIN 429 442 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 443 463 12 (POTENTIAL).
FT DOMAIN 464 491 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 64 64 A -> V (IN REF. 6).
SQ SEQUENCE 491 AA; 53608 MW; 2AF1AF9756C0B722 CRC64;
Query Match 10.7%; Score 434; DB 1; Length 491;
Best Local Similarity 19.5%; Pred. No. 7.5e-21;
Matches 152; Conservative 109; Mismatches 175; Indels 344; Gaps 17
QY 62 VAIVASIGNLLQGWDNATIAAVALYIKKEF--QLONEPTVEGLI---VSMSLIGATIVT 115
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 13 ITLVATLIGLLFGYDTAVIGTVESLNTVFVAPQLNSESANSLIGFCVASALIGCIIGG 72
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 TRSGPLSDSIGRRPMLILSSILYFFSGLIMLW-----SPN-----VYV-----LLLA 157
Db 73 ALGVCYCSNRFGRDSIKIAAVLFFISGVGSAMPGLGFTSINPDNTVPYLAGVPEFVIY 132
QY 158 RPYDGFICGLAVTLVPLVYISEIARSEIRGLLNTLPQFSSGGMFLSYCMVFGMSLS---- 213
Db 133 RIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSPNQFAIIFGQLLVYCVNYFIARSGDAS 192
QY 214 --PSPDKRIMLGVLAIPSLFFFGGLTIFYLPESPRLVSKGRMAEAKVLOKLRGKDDVSG 271
Db 193 WINTDGWRMYFASECIPALFL-MLLYTVESPRLMSRGKQQAEGILRKIMGNT---- 247

```

QY	272	ELSLLEGEVGGDTSIEEYIGPATEAADLVTDDKEQITLYGPEEGQSWIARPSKCP	331
Db	248	-----LATQAVQEI-----	256
QY	332	IMIGSVLSLASRHGSMVNQSVPLMDPIVTLFGSVHNMPPQAGGSMRSTLFPNFGSMFSVT	391
Db	257	-----	256
QY	392	DQAKNEQWDEENLHRDDEEYASDAGAGDYEDNLH3PLLSRQATGAEGKDIVHHGRGSA	451
Db	257	-KHSL-----	260
QY	452	LSNRQOTLLGEGGDVSTDIGGQWLAMWKSEKEGENRKEGFEKRVYLHQEYDGSRR	511
Db	261	-----DHGRKTGG--RLLMFGVGV-----	277
QY	512	GSIVSLPGCGDVFEGSEFVHAALVSQSALFSKGLAEPMSDAAAMVHPSEVAAKGSRWKD	571
Db	278	-----	277
QY	572	LFEPGVRALLVGVGIQILOQFAGINGVLYTTPQILLQAGVAVILSKFGLSSASISILIS	631
Db	278	-----IVIGVMLSIPOQFVGINVLYYADEVFKTLG-----ASTDIALLOT	318
QY	632	SLTTLMLPCIGFAMLLMDLSGRFL-LIGTIPILLASLVILVSNLIDLGTL---AHA	686
Db	319	IIVGVINLFTVLAIMTVDKFGRKPLQIGALGMAT-----GMFSLGTAFYTOAPG	369
QY	687	LLSTVSVIVYFCCFVMGFGPIPNILCAEIPTRVRGLCIAICAFTEWIGDIIYTVYSLPVM	746
Db	370	IYALLSMLFYAAAFAMSWGPPVCWVLLSEIFPNAIRKALAIAYAAQWLANYFVSWTFPMM	429
QY	747	LNAIGLAVFS-----IYAVVCLISFVFPVLKVPBTKGMPLEVITTEFPAVGAKQAATAKA	800
Db	430	DKNSWLVAHFHNGFSYWIYGCMMGVLAALFMWKFPVPHTKGKTLEELLEALWEPETKKTQOTA	489

RESULT 9	ITR2_SCHPO	STANDARD;	PRT; 557 AA.
ID	ITR2_SCHPO		
AC	P87110; P78501;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Myo-inositol transporter 2.		
GN	ITR2 OR SPAC20G8.03.		
OS	Schizosaccharomyces pombe (Fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OX	Schizosaccharomycetes.		
NCBI_TaxID=4896;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=968 h90;		
RX	MEDLINE=98228265; PubMed=9560432;		
RA	Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,		
RA	Rusu M., Poltelea M., Edenhartter L., Schweingruber M.E.;		
RT	"Exogenous inositol and genes responsible for inositol transport are		
RL	required for mating and sporulation in Shizosaccharomyces pombe.";		
RL	Curr. Genet. 33:255-261(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RX	MEDLINE=21848401; PubMed=11859360;		
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne N., Lyne R., Stewart A.,		
RA	Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,		
RA	Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,		
RA	Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,		
RA	Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,		
RA	Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,		
RA	James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,		
RA	Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,		
RA	Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,		

```

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Wellstead I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussey D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RA Nature 415:871-880(2002).
RL [3]
RN
RP SEQUENCE OF 166-541 FROM N.A.
RC STRAIN=PR745;
RC MEDLINE=98162722; Pubmed=9501991;
RX Yoshioaka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cdna.";
RT DNA Res. 4:363-369(1997).
RL
CC -1- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL, X99105; -, NOT ANNOTATED_CDS.
CC
CC EMBL, D95334; CAB08597.1; -.
CC EMBL, D89252; BAA13913.1; -.
CC PIR, T38125; T38125.
CC
CC GeneDB_SPombe; SPAC20G8.03; -.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transpt.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGRTRANSPORT.
CC TIGRfams; TIGR00879; SP; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
CC Transmembrane; Sugar transport.
CC
CC DOMAIN 1 76 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 77 97 1 (POTENTIAL).
CC DOMAIN 98 99 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 100 120 2 (POTENTIAL).
CC DOMAIN 121 123 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 124 144 3 (POTENTIAL).
CC DOMAIN 145 157 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 158 178 4 (POTENTIAL).
CC DOMAIN 179 180 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 181 201 5 (POTENTIAL).
CC DOMAIN 202 209 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 210 230 6 (POTENTIAL).
CC DOMAIN 231 240 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 241 261 7 (POTENTIAL).
CC DOMAIN 262 367 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 368 388 8 (POTENTIAL).
CC DOMAIN 389 396 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 397 417 9 (POTENTIAL).
CC DOMAIN 418 432 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 433 453 10 (POTENTIAL).
CC DOMAIN 454 468 CYTOPLASMIC (POTENTIAL).

```



```

DB      5  VMLCAALGGFLFGYDTGVINAAALFQMKDHFGFSEHSMQYALIVAIAIAGAFAVGAFISGFI 64
QY      122 SDSIGRRPMLILSSILYFFSGILMLWSPNVYLLARFVDFGIGLAVTLVPLYISEIAP 181
DB      65  SAAFGRRPCIAVADALFVIGSVLMGAAPNVEVLVSRVYGLAIGISSATIPVYLAEVTS 124
QY      182 SEIRGLNLTLPQFSSGGMFLS--YCMVFEGMSLSPSPWRIMLGVALPSLFFFGJLIFY 239
DB      125 PKHGGATIVLNNLFLTGGQFVAAGFTALMVFTSKNIGWVAIGIGALPAVVGQAFCLLFF 184
QY      240 LPESPRWLVSXGRMAEAKVYLOKLRGKDDVSGELSLLEGLEVGDDTSIEEYITGPATEA 299
DB      185 LPESPRWLSSKHADRAKAV----- 204
QY      300 ADDLVTDGKEQITLYGPEEGOSIARPSKGPIMLSVLSLASRHGSMVNQSVPLMDPIV 359
DB      205 -----ADKFEVDLCFEQEG----- 218
QY      360 TLFGSVHNMPOAGGSMRSTLPNFGSMFSVTDQHAKEQWDEENLHRDDEEYASDAGG 419
DB      219 -----DELP-----SVRI----- 226
QY      420 DYEDNLHSPILSRQATGAEGKDIVHGHRSALSMRQRTLLGEGGDVSTDIGGQOLA 479
DB      227 DYR-----PLMARD----- 235
QY      480 WKHSEKEGENGRKEGFKRVYLHQEGVPGSRGSIYSLPGGDVEGSEFVHAALVSQS 539
DB      236 ----- 235
QY      540 ALFSKGLAEPMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVCIQILQOPAGINGV 599
DB      236 -----MRFRVLLSSGLQIIQFSGINTI 258
QY      600 LYTPQILBQAGVAVILSKFGLSSASASILSSLTTLMLPCIGFAMLLMDLGGRRFLL 659
DB      259 MYSSVILYDAGFRDAIMPVLSIPLA--FMNALFTAV-----AIFTVDRFGRRL 309
QY      660 GTIPILIASLVILVSNLIDL-----GTLAHLSTVSVIVYFCCFVMGEGPIP 708
DB      310 ISV---FGCLVLLVVAIIGFFIGTRISYSVGGGLFLALLA----VFALAYAPGICIP 361
QY      709 NILCAIEPTVRVRLCAICAFTEWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFV 768
DB      362 WVMGEIFPTLRTSASAVATMANWGANVLVSQVFPILMGALIGVGFTTISGLMALGCI 421
QY      769 FVFLKVPETKGMPLLEVITEFF 789
DB      422 FVYFPAVETKGLTLEQIDNMF 442

RESULT 11
ARAE_ECOLI
ID      ARAE_ECOLI          STANDARD;          PRT;          472 AA.
AC      P09830; Q46937;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Arabinose-proton symporter (Arabinose transporter).
GN      ARAE OR B2841 OR Z4161 OR ECS3698.
OS      Escherichia coli, and
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=562, 83334;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87115869; PubMed=3543693;
RA      Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA      Henderson P.J.F.;
RT      "Mammalian and bacterial sugar transport proteins are homologous.";
RL      Nature 325:641-643(1987).
RN      [2]

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / JM2433;
RX      MEDLINE=88228015; PubMed=2836407;
RA      Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
RT      "The cloning, DNA sequence, and overexpression of the gene arae
RT      coding for arabinose-proton symport in Escherichia coli K12.";
RL      J. Biol. Chem. 263:8003-8010(1988).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blattner F.R.;
RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL      Nature 409:529-533(2001).
RN      [5]
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / RIMD 0509952;
RX      MEDLINE=21156231; PubMed=11258796;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Ikeda T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA      Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT      "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT      O157:H7 and genomic comparison with a laboratory strain K-12.";
RL      DNA Res. 8:11-22(2001).
RN      [6]
RP      PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
RX      MEDLINE=84114868; PubMed=6319708;
RA      Stoner C., Schleif R.F.;
RT      "The arae low affinity L-arabinose transport promoter. Cloning,
RT      sequence, transcription start site and DNA binding sites of
RT      regulatory proteins.";
RL      J. Mol. Biol. 171:369-381(1983).
CC      -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC      THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC      -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J03732; AAA23469.1; -.
DR      EMBL; X00272; CAA25075.1; -.
DR      EMBL; U29581; AAB40488.1; -.
DR      EMBL; AE000368; AAC75880.1; -.
DR      EMBL; AE005513; AAG57953.1; -.
DR      EMBL; AP002563; BAB37121.1; -.
DR      PIR; B26430; B26430.
DR      PIR; B91091; B91091.
DR      PIR; B85936; B85936.
DR      EcoGene; EG10056; arae.
DR      InterPro; IPR007114; MPS.
DR      InterPro; IPR005828; Sub_transporter.

```

```
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
Complete proteome.
FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 30 50 1 (POTENTIAL).
FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 64 84 2 (POTENTIAL).
FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 92 112 3 (POTENTIAL).
FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 115 135 4 (POTENTIAL).
FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 5 (POTENTIAL).
FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 179 199 6 (POTENTIAL).
FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 258 278 7 (POTENTIAL).
FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 298 318 8 (POTENTIAL).
FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 326 346 9 (POTENTIAL).
FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 362 382 10 (POTENTIAL).
FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 405 425 11 (POTENTIAL).
FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 428 446 12 (POTENTIAL).
FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
SQ SEQUENCE 472 AA; 51684 MW; 411990A41D44393 CRC64;

Query Match 10.5%; Score 429.5; DB 1; Length 472;
Best Local Similarity 20.6%; Pred. No. 1.4e-20;
Matches 154; Conservative 100; Mismatches 194; Indels 299; Gaps 13;

QY 52 LEDKMSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQLONEPTVEGLIVSNLIGA 111
DB 15 LRDTRRMNFVSAVAAGLLEGLDIGIVAGALPFIIDHFLVTSR--LQEWVSSMWLGA 72
QY 112 TIVTFSGPLSDSIGRPMILISSILYFSGLIMLWSPNVYLLARFVDGFGIGLAVTL 171
DB 73 AIGALFNGWLSFRLGRKYSIMAGAILFVLGSIGSAFATSEMLIARVVLGIAVGIASYT 132
QY 172 VPLYISEIAPSEIRGLNLTLPQFSSGGMFLSYCMVFGMSLSPSPDWIRIMLGVALPSLF 231
DB 133 APLYISEMASENVRGKMTSMYQMLWTLGIVLAFL--LSDTAFSYSGNMRAMLGVLAIPVL 190
QY 232 FFGLTIFYLPESPRLVSKGRMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIEY 291
DB 191 LILLVVF-LPNSPRWLAEKGRHIEAEVLRMLR-----DTS----- 225
QY 292 IIGPATEADDLVTGDKEQITLYGPEEGQSWIARPSKPIMLGSLSLASRHGSMVNS 351
DB 226 -----EKAREEL-----NEIRES 238
QY 352 VPLMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSVTQHAKNQOWDEENLRDDEE 411
DB 239 LKLKQ----- 243
QY 412 YASDAGAGDYEDNLHSPLSRQATGAEGKDIVHHGHRGSALSMRQTLGEGGDVSTSD 471
DB 244 ----- 243
QY 472 IGGGWQLAMKMSKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGDFEGSEFVH 531
DB 244 --GGW----- 246
```

```
QY 532 AALVQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEEGVRALLVGVGIQLQ 591
DB 247 -----ALFK-----INRNVRAVFLGMLLQAMQ 269
QY 592 QFAGINGLVYTPQILEQAGVAVILSKFGLSSASASILSSLTLLMPCIGFAMLLMDL 651
DB 270 QFTGMNIIMYAPRIEFKMA-----GFTTEQQMIALVLVGLTFMFATPIAVFTVDK 320
QY 652 SGR-FLLAGTIPILIASLIVLVNSLIDLTALHALLSTVSIVVFCC---FVMGFGPI 707
DB 321 AGRKPAKIGFSVALGTLVLGYCLMOPDNGT-ASSGLSWLSVGMTMCTAGYAMSAAPV 379
QY 708 PNILCAEIPTRVRGLCIACAFTFWIGDITVYSLPVMINAIGLAVGFSIYAVCLISF 767
DB 380 VWILCSEIQPLKCRDFGTCSTTNWVSNMIGATFLLTLDISGAAGTFWLYTALNIAFV 439
QY 768 VFVFLKVPETKGMPELVITEFFAVGAK 794
DB 440 GITFWLIPETKAVTLEHIERKLMAGEK 466

RESULT 12
ITR1_YEAST
ID ITR1_YEAST STANDARD; PRT; 584 AA.
AC P30605;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 1.
GN ITR1 OR YDR497C OR D9719.3.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250431; PubMed=2040626;
RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT "Isolation and characterization of two distinct myo-inositol
transporter genes of Saccharomyces cerevisiae.";
RL J. Biol. Chem. 266:11184-11191(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oelfner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; D90352; BAAL4366.1; -.
DB EMBL; U33057; AAB64939.1; -.
DB PIR; S69555; S69555.
DB SGD; S0002905; ITR1.
DB GO; GO:0005365; F:myo-inositol transporter activity; IMP.
DB GO; GO:0015798; P:myo-inositol transport; IMP.
DB InterPro; IPR007114; MFS.
DB InterPro; IPR005828; Sub_transporter.
DB InterPro; IPR005829; Sug_transporter.
DB InterPro; IPR003663; Sugar_transpt.
DB Pfam; PF00083; sugar_tr; 1.
```


DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 81 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 182 102 1 (POTENTIAL).
FT DOMAIN 103 129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 130 150 2 (POTENTIAL).
FT DOMAIN 151 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 184 3 (POTENTIAL).
FT DOMAIN 185 186 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 187 207 4 (POTENTIAL).
FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 5 (POTENTIAL).
FT DOMAIN 237 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 370 7 (POTENTIAL).
FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 8 (POTENTIAL).
FT DOMAIN 398 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 9 (POTENTIAL).
FT DOMAIN 422 441 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 442 462 10 (POTENTIAL).
FT DOMAIN 463 486 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 507 11 (POTENTIAL).
FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 12 (POTENTIAL).
FT DOMAIN 532 584 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 43 44 TL -> HI (IN REF. 1).
SQ SEQUENCE 584 AA; 63569 MW; 42543E30A102DC65 CRC64;

Query Match 10.3%; Score 421.5; DB 1; Length 584;
Best Local Similarity 19.1%; Pred. No. 6.1e-20;
Matches 150; Conservative 121; Mismatches 195; Indels 321; Gaps 19;

QY 44 EPTVSD-----ILEDKMGAVL-VAIVASIGNLQGMNATIAAVALYIKKEFQONE 96
DB 64 KPVNDEDDTSVMITFNQSLSPFIITLTFVASISGFMGYDTGYISSALISIGTD--LDHK 121
QY 97 PVEG--LIVSMSLGATIVTFESGPLSDSIGRRPMLISSILYFSGLIMLMSPNVYV 153
DB 122 VLTGEKEIVTAATSLGALITSIFAGTAADIFGRKRCMLGNSLMFVIGAILQVSAHTFQ 181
QY 154 LLLARFVDGFGIGLAVTLVPLYISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLS 213
DB 182 MAVGRLLIMFGVGIGSLIAPLFISEIAPKMRGLTVNSLMLTGGQLVAYCGGALNY- 240
QY 214 PSPDWRIMGLVLAIPSLFFEGITIFYLPESPRMLVSKGMAEAKVLOKLRKDVSGEL 273
DB 241 VNNGWRILVGLSLIPTAVQF-TCLCFLPDTPRYYVMKGLARATEVLK--RSYDTSEEI 297
QY 274 SLLLEGLVGGDTSIEEYIIGPATEADDLVTDGKEQITLLYGPEEGQSWIARPSKPIIM 333
DB 298 -----IERKV-----ELVT----- 307
QY 334 LGSVLSLASRHGSMVNOVPLMDPIVTLFGSVHENVPOAGGSMRSTLFPNFGMSRSTVDQ 393
DB 308 -----LNQSTP----- 313
QY 394 HAKNEQWDEENLHRDEEYASDAGGDYEDNLHSPILSRQATGAEKDIYHHGRGSALS 453
DB 314 -----GKNV----- 317
QY 454 MRQOTLLGEGGDVSSTDIGGQWLAKWSEKEGENKRGKRVYLLHQEGVPSRRGS 513
DB 318 ----- 317
QY 514 IVSLPGGDFEGSEFYHAALVSQALFSKGLAEFRMSDAMVHPSEVAAGSRWKDLF 573
DB 318 -----PEKV-----WNTIK 326

QY 574 E----PGVRALLVGVGIQIQFAGINGVLYTPQILEQAGVAVILSKRGISSASATL 629
DB 327 ELHTVPSNLRALIGCGLOAIQOFTGMNSLMYFSGTIFETVGFK-----NSSAVSI 378
QY 630 ISSLTLLMPCIGFAMLLMDLSGRRELLGTPILIASVL-----VSN 676
DB 379 VSGTNFIPTL---VAFFSIDKIGRTILLIGPMTMLVVCISIAFHILGKFDGAVAV 434
QY 677 LIDGLTAHALSTVSIVYFCCFWMGFPIPNILCAEIPTRVRGLCIACAFTWIGD 736
DB 435 VVSSGSSWGIIVIFIVEFAFYALGIGYD-WQOSELPQNVRGIGTSYATATNWAGS 493
QY 737 IIVTYSLPVMLNAGLAGVESIYAVVCLISFVFLKVPETKMPLE---VTEFFAVG 792
DB 494 LVIASTFLTMLQNTIPAGTFAFAGLSCLSTFCYFCYPELSGLELEEVQTLKGFNIK 553
QY 793 AKQAAK 799
DB 554 ASKALAK 560

RESULT 13

MYCT_HUMAN STANDARD; PRT; 629 AA.
ID MYCT_HUMAN
AC Q96QE2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proton myo-inositol co-transporter (Hmlt).
GN SLC2A13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21391710; PubMed=11500374;
RA Uldry M., Ibberson M., Horisberger J.-D., Chatton J.-Y.,
RA Riederer B.M., Thorens B.;
RT "Identification of a mammalian H(+)-myo-inositol symporter expressed
RT predominantly in the brain."
RL EMBO J. 20:4467-4477(2001).
CC -!- FUNCTION: H(+)-myo-inositol co-transporter. Can also transport
CC -!- related stereoisomers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed in the brain.
CC -!- PTM: Glycosylated.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ315644; CAC5116.1; -.
CC Genew; HGNC:15956; SLC2A13.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 1 (POTENTIAL).
FT DOMAIN 79 106 EXTRACELLULAR (POTENTIAL).

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X98622; CAA67211.1; -.
DR EMBL; Z98530; CAB11061.1; -.
DR EMBL; Z99532; CAB16718.1; -.
DR PIR; T43400; T43400.
DR Genedb_Spombe; SPAC4F8.15; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRANSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 87 107 1 (POTENTIAL).
FT DOMAIN 108 129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 130 150 2 (POTENTIAL).
FT DOMAIN 151 156 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 157 177 3 (POTENTIAL).
FT DOMAIN 178 186 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 187 207 4 (POTENTIAL).
FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 5 (POTENTIAL).
FT DOMAIN 237 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 370 7 (POTENTIAL).
FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 8 (POTENTIAL).
FT DOMAIN 398 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 9 (POTENTIAL).
FT DOMAIN 422 441 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 442 462 10 (POTENTIAL).
FT DOMAIN 463 486 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 507 11 (POTENTIAL).
FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 12 (POTENTIAL).
FT DOMAIN 532 575 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 575 AA; 62757 MW; 3B7C5EPF86C596AE CRC64;

Query Match 9.9%; Score 404.5; DB 1; Length 575;
Best Local Similarity 19.4%; Pred. No. 7.5e-19;
Matches 153; Conservative 104; Mismatches 192; Indels 341; Gaps 17;

QY 38 ASGSGEPVTSDDILEDK--MSGAVLVAIVASIGNLLOGWDNATIAAALVYI-----87
Db 64 SNISLSEPHALNDTVEDQPSKMWVLAFAAGIGLLFGYDTGVISGALVIGTSLGHE 123
QY 88 -----KKEFQLONEPTVEGLIVSMGLIGATIVTTFSGPLSDSIGRRPMLILSSILYFFSGL 143
Db 124 LTNGKEF-----ITSATSLGALLGIIAGALADFGRRKPVIAIASIILYVGS 172
QY 144 IMLSPNVYVLLARFVDFGIGLAVTLVPLYISEIAPSEIRGLINTLPQFSGSGGMFLS 203
Db 173 VQVTAHHLHMIVGRFVIGWGVGLASLIIPLYLSEIAPSKIRGLVITIVLLITAGQVIA 232
QY 204 YCMVFGMSLS---PSPDWIRIMGLVLAIPSLFPFGLTIFYLPSBPRLVSKGRMAEAKVYL 260
Db 233 Y---GIDTAFEHVHNGWRMVGILANVPFAFQLFILI-WLPESPRLLVKERSQEAAYNTL 287
QY 261 QKLRGKDDVSGELSLLEGLEVGDTISIEYIIGPATEAADLVTDGKEQITLYGPEEG 320
Db 288 ARI-----290

QY 321 QSWIARPSKPIMLGSVLSASRHGSMVNSVPLMDPIVTLFGSVHENNPQAGSMRSTL 380
Db 291 -----YPTAHP-----296
QY 381 PPNFGSMFSVTDQAKNEQWDEENLHRDEDEYASDGAGDYEDNLHSPILSRQATGAEGK 440
Db 297 -----YEIK-----300
QY 441 DIVHGHGHSALSMRRQTLLEGGGDVSTDIGGQWLAWKWKSEKEGNGRKEGFKRVY 500
Db 301 -----TKLY 304
QY 501 LHQGVPGSRGSIYSLPGGDVFEFGSEFVHAALVSQSALFSKGLAEPMSDAMVHPS 560
Db 305 LIQGV-----RDPF-----314
QY 561 EVAAKGSRWKDL-----FEPGVRALLVGVIQIILQOFAGINGVLYTTPQILLEQAGV 612
Db 315 ----SGSRWQKIYKTFKELYFNPSNFRALLIACGLQAMQOLSGFNSLMYFSSITFEVVG- 369
QY 613 AVILSKFGLSSASILISSLTLLMPCIGFAMLLMDLSGRFLLGTFPIILIASLVIL 672
Db 370 -----FNNPTATGLITAA--TNFVFTIYAGVI--DFGRRILLLLTYGMIALIVC 418
QY 673 VVS-----NIDLTALHALSTVSIVYFCCFVMGFGPIPNILCAEIPTRVRG 722
Db 419 AVAFHFLPKDENGNYTSGQSNAMAIIVLISMIYVASYASGLGNLP-WQOSELPMSVRG 477
QY 723 LCIAICAFTFWIGDIIVTVSLPVMNLNAGIAGVFSIYAVVCLISFVFVFLKVBETKGMPL 782
Db 478 LGTGMSTAVNWAGNIGIGASFLTLMSEITPTGTFALYGLCFLGWLGAFCYFIDLTDYTI 537
QY 783 EVITTEFAVG 792
Db 538 EIIGELLKHG 547

Search completed: January 5, 2004, 18:56:02
Job time : 36 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:09:20 ; Search time 95 Seconds
(without alignments)
2173.075 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPLDDR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL_23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriaph:*
 - 17: sp_archaeap:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2834	69.5	740	10 Q9FRK0	Q9frk0 oryza sativ
2	2808	68.9	743	10 Q8GT52	Q8gt52 hordeum vul
3	2654	65.1	739	10 Q8LPG8	Q8lpg8 arabidopsis
4	2584	63.4	729	10 Q65497	Q65497 arabidopsis
5	2278	55.9	734	10 Q9SYQ3	Q9syq3 arabidopsis
6	2275	55.8	734	10 Q9LME7	Q9lme7 arabidopsis
7	2245	55.1	729	10 Q9SD00	Q9sd00 arabidopsis
8	2191.5	53.8	753	10 Q8GT51	Q8gt51 hordeum vul
9	2183	53.6	714	10 Q96290	Q96290 arabidopsis
10	1763	43.3	714	10 Q8H887	Q8h887 oryza sativ
11	1763	43.3	723	10 Q9SNK7	Q9snk7 oryza sativ
12	1270.5	31.2	746	10 Q8LJX7	Q8ljx7 sorghum bic
13	1050	25.8	643	10 Q8H6J2	Q8h6j2 zea mays (m
14	621.5	15.3	664	10 Q9XEVI	Q9xevi oryza sativ
15	587.5	14.4	457	16 P96742	P96742 bacillus su
16	536	13.2	580	10 Q9C757	Q9c757 arabidopsis

17	535.5	13.1	581	10 Q9LKH2	Q9lkh2 mesembryant
18	524.5	12.9	580	10 Q9ZQP6	Q9zqp6 arabidopsis
19	520.5	12.8	582	10 Q23492	Q23492 arabidopsis
20	519	12.7	581	10 Q9LKH1	Q9lkh1 mesembryant
21	518.5	12.7	469	16 Q97JE7	Q97je7 clostridium
22	512.5	12.6	577	10 Q945E5	Q945e5 oryza sativ
23	505	12.4	473	16 Q34718	Q34718 bacillus su
24	500	12.3	509	10 Q8VZR6	Q8vzr6 arabidopsis
25	498	12.2	463	16 Q8EMP5	Q8emp5 oceanobacil
26	494	12.1	470	10 Q9LLE0	Q9lle0 solanum tub
27	493.5	12.1	508	10 Q9ZNS0	Q9zns0 arabidopsis
28	493	12.1	493	10 Q23213	Q23213 arabidopsis
29	492	12.1	467	16 Q8CQA7	Q8cqa7 staphylococ
30	492	12.1	534	10 Q9LLE1	Q9lle1 nicotiana t
31	489	12.0	521	10 Q22848	Q22848 arabidopsis
32	488.5	12.0	502	10 Q9FRP7	Q9frp7 oryza sativ
33	488	12.0	545	10 P93076	P93076 beta vulgar
34	488	12.0	549	10 P93075	P93075 beta vulgar
35	485.5	11.9	539	10 Q9LS92	Q9ls92 arabidopsis
36	481	11.8	513	10 Q9FQX3	Q9fqx3 apium grave
37	474	11.6	544	10 Q93WT7	Q93wt7 olea europa
38	474	11.6	547	10 Q9SKT9	Q9skt9 arabidopsis
39	469.5	11.5	515	10 Q9LID8	Q9lid8 arabidopsis
40	468	11.5	551	10 Q9LLE2	Q9lle2 spinacia ol
41	467	11.5	546	10 Q93Z41	Q93z41 arabidopsis
42	465	11.4	419	16 P94493	P94493 bacillus su
43	465	11.4	506	10 Q8W2W8	Q8w2w8 oryza sativ
44	465	11.4	542	10 Q9LID9	Q9lid9 zea mays (m
45	465	11.4	574	10 Q9AUM9	Q9aum9 oryza sativ

ALIGNMENTS

RESULT 1	ID	Q9FRK0	PRELIMINARY;	PRT;	740 AA.
AC	Q9FRK0;	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)				
DT	01-MAR-2003 (TREMBLrel. 23, last annotation update)				
DE	Putative sugar transporter.				
GN	OSJNB0064P21.3.				
OS	Oryza sativa (Rice).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
OC	Eurharoidae; Oryzae; Oryza.				
OX	NCBI_TaxID=4530;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Nipponbare;				
RA	Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,				
RA	Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,				
RA	Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,				
RA	Quackenbush J., White O., Salzberg S.L., Fraser C.M.;				
RT	"Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence.";				
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.				
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).				
DR	EMBL; AC073166; AAG46115.1; -				
DR	Gramene; Q9FRK0; -				
DR	InterPro; IPR007114; MFS.				
DR	InterPro; IPR005828; Sub_transporter.				
DR	InterPro; IPR003663; Sugar_transp.				
DR	InterPro; IPR005829; Sug_transporter.				
DR	Pfam; PF00083; sugar_tr; 1.				
DR	PRINTS; PR00171; SUGRTNSPORT.				
DR	PROSITE; PS00850; MFS; 1.				
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.				
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.				
KW	Sugar transport; Transmembrane; Transport.				
SO	SEQUENCE 740 AA; 78899 MW; D4C003E0A4980C46 CRC64;				

DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE AT4G35300/F23E12.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AY094465; AAM19835.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 2.
DR PRINTS; PR00171; SUGRTRANSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 739 AA; 79725 MW; 5D67C6D2DE7563FD CRC64;

Query Match 65.1%; Score 2654; DB 10; Length 739;
Best Local Similarity 70.0%; Pred. No. 5.1e-172;
Matches 521; Conservative 90; Mismatches 123; Indels 10; Gaps 7;
QY 56 MSGAVLVAIVASIGNLLOGMDNATIAAAVLYIKKEFOLNEPTVEGLIVSMISLIGATIVT 115
DB 1 MSGAVLVAIAAAGNLLQGMNATIGAVLYIKKEFNESNPSEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLISLILYFSGLIMLWSPNVYVLLARFVDGFGIGLAVTVPLY 175
DB 61 TCSGGVADWLGRRPMLISLILYFVGSVLMWSPNVYVLLIGRLLDGFGVGLVTVPLY 120
QY 176 ISEIAPSEIRGLNLTLPQSSGGMFLSYCMVFGLSPSPDWIRMLGVLAIPSLFFEGL 235
DB 121 ISETAPEIRGLNLTLPQFTSGGMFLSYCMVFGLSPSPDWIRMLGVLAIPSLFFEFL 180
QY 236 TIFYLPESPRWLVSKGMAEAKVYLQKRGKDVSGELSLLEGLVEGDTSEIETIGP 295
DB 181 TVFFLPESPRWLVSKGMAEAKVYLQKRGREDVSGEMALLVEGLIGGETTIEETIGP 240
QY 296 ATEAADLVTDGDKQITLYGPEGQSWIARPSKPIMLGSVLSLASRHGS-MVNGSVPL 354
DB 241 ADEVTDHDIADVDDQIKLYAEGLSWARPVKG---GSTMSVLSRHGSTMSRQSL 296
QY 355 MDPVITLFGSVHNMPOAGGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDEETAS 414
DB 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPFHGSMFVSGNDRHEDWDEENLVGEEDYPS 355
QY 415 DGAGGDYEDNLHSPLLSRQATGAEGKDIVHGHRSALSMRQTLLEGEGDVSTIDIG 474
DB 356 D-HGDSEDDLHSPILSRQTSME-KDMPHTAHGTLSTFRHGSQVQAGGEGAGSMGIG 413
QY 475 GWQLAWKWEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDFEESSEFYHAA 534
DB 414 GWQYAWKWTREDESGQKEGFKRITYLHQEGFPGSRGSIIVSLPGDGTGE-ADFYQSA 472
QY 535 LVSQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQQA 594
DB 473 LVSQPALYSKDLLKEHTTIGPAMVHPSE-TTKGSIWHDLHDPGVKRALVVGVLQILQQS 531

QY 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTTLLMPCIGFAMLLMDISGR 654
DB 532 GINGVLYTTPQILEQAGVIGILSNMGISSSSASLISALTTFVMLPAIVAMRLMDISGR 591
QY 655 RELLGTIPILIASLVILVSNLIDGLTAHALISTVSIVIVFCCFWMGFGPIINILCAE 714
DB 592 RTLLLTIPILIASLVILVSNLIDGLTAHALISTVSIVIVFCCFWMGFGPAPNIIICSE 651
QY 715 IFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFLKV 774
DB 652 IFPTRVRGICIAICAFTEWICDIIVTYSLPVLLKSLIGLAVFGMYAIVCCISWFEVFIKV 711
QY 775 PETKGMPLVITEFPFVGAQAQAA 798
DB 712 PETKGMPLVITEFPFVGARQAQA 735

RESULT 4
ID 065497 PRELIMINARY; PRT; 729 AA.
AC 065497;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Putative sugar transporter protein (Monosaccharide sensing protein 2).
GN F23E12.140 OR AT4G35300 OR MSSP2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Hohelsel J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.F.X.,
RA Schueller C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Stamme C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
RT "A novel family of plant monosaccharide transporters is involved in a
new type of eukaryotic sugar sensing."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AL022604; CAI18739.1; -.
DR EMBL; AL161587; CAB80247.1; -.
DR EMBL; AJ532570; CAD58692.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRANSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 729 AA; 78524 MW; 81A35DDF8725411E CRC64;

Query Match 63.4%; Score 2584; DB 10; Length 729;
Best Local Similarity 68.8%; Pred. No. 2.9e-167;

Matches 512;		Conservative 89;	Mismatches 123;	Indels 20;	Gaps 8;
QY	56	MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQOLNEPTVEGLIVSMSLIGATIVT	115		
Db	1	MSGAVLVAIAAAGVNLQGWDNATIAAGVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT	60		
QY	116	TFSGPLSDSIGRRPMLILSSILYFFSGLIMLMSPNVYVLLARFVDGFGIGLAATVPLV	175		
Db	61	TCSGGVADWLGRRPMLILSSILYFVGSVLWMSPNVYVLLGRLLDGFVGLVTLVPIY	120		
QY	176	ISEIAPSEIRGLNTLPQFSGSGGMFLSYCMVFEGMSLSPSPDWIRIMLGVALIPSLFFPGL	235		
Db	121	ISETAPPEIRGLNTLPQFTSGGGMFLSYCMVFEGM3LMPSPSWRLMLGVLFIPLVFFFL	180		
QY	236	TIFYLPESPRWLVSKGMAEAKVLOKLRKDDVSGELSLLEGLVEGDTISIEYIIGP	295		
Db	181	TVEFLPESPRWLVSKGRLMAKRVLOLRGREDVSGEMALLVEGLIGGETTIEEYIIGP	240		
QY	296	ATEAADLVTDGKEQITLYGPEGQSWIARPSKGPIMLGSVLSLARHGS-MVNGSVPL	354		
Db	241	ADEVTDHDHIAVDKQIKLYGAEGLSWVARPVKG---GSTMSVLSRHGSTMSRRGSL	296		
QY	355	MDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSVTQDAKNEQWDEENLHRDEEYAS	414		
Db	297	IDPLVTLFGSVHEKMPDT-GSMRSALFPHFGSMFSVGNDQRHEDWDEENLVGEDEYPS	355		
QY	415	DGAGGDYEDNLHSPILSRQATGAEGKDIVHGHRSALSRRQTLLEGGDGVSTIDIG	474		
Db	356	D-HGDDSEDDLHSPILSRQTSME-KDMPHTAHGT,STRHGSQVQGAQEGAGSGMIGG	413		
QY	475	GWQLAWKMKSEKEGNGRKEGKRVYLHQEGVPGSRGSIIVSLPGGDFEGSEFPHAA	534		
Db	414	GWQVAMKWKTEREDESQKE-----EGFPGRSGSIIVSLPGDGTGE-ADFYQASA	462		
QY	535	LVSQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVIQIIQFA	594		
Db	463	LVSQPALYSKDLKEHTIGPAMVHPSE-TTKGSIWHDLHPGVKRALVVGVLQIIQFS	521		
QY	595	GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTTLLMLPCIGFAMLLMDLSGR	654		
Db	522	GINGVLYTTPQILEQAGVAVILSNMGISSSSASASILISALTFFVMLPAIVAMRLMDLSGR	581		
QY	655	RELLLTIPILIASLVILVNSNLDLGLAHALLSTVSIVTYFCCFVMGFGPIPNILCAE	714		
Db	582	RTLTLTIPILIASLVILVNSLVHANSIVHAVLSTVSIVLYFCFVMGFGPAPNIIQSE	641		
QY	715	IFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLKV	774		
Db	642	IFPTRVRGICIAICALTFWICDIIIVTYSLPVLKLSIGLAGVFGMYAIVCCISWVFVIKV	701		
QY	775	PETKGMPLLEVITEFFAVGAKQAAA 798			
Db	702	PETKGMPLLEVITEFFSVGARQAEA 725			
RESULT 5					
Q9SYQ3		PRELIMINARY; PRT; 734 AA.			
AC Q9SYQ3:					
DT 01-MAY-2000 (Tremblrel. 13, Created)					
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)					
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)					
DE Sugar transporter.					
GN F9H16.18.					
OS Arabidopsis thaliana (Mouse-ear cress).					
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;					
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.					
OX NCBI_TaxID=3702;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=cv. Columbia;					
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,					

RA	Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,	
RA	Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,	
RA	Lucos S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,	
RA	Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;	
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.	
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).	
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.	
DR	EMBL; AC007369; AAD30608.1; -.	
DR	InterPro; IPR007114; MFS.	
DR	InterPro; IPR005828; Sub_transporter.	
DR	InterPro; IPR003663; Sugar_transpt.	
DR	InterPro; IPR005829; Sug_transporter.	
DR	Pfam; PF00083; sugar_tr; 1.	
DR	PRINTS; PR00171; SUGRTRANSPORT.	
DR	PROSITE; PS50850; MFS; 1.	
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.	
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	
KW	Sugar transporter; Transmembrane; Transport.	
SQ	SEQUENCE 734 AA; 79426 MW; AB027864F850259C CRC64;	
Query Match 55.9%; Score 2278; DB 10; Length 734;		
Best Local Similarity 61.2%; Pred. No. 2e-146;		
Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;		
QY	56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQOLNEPT-VEGLIVSMSLIGATIV	114
Db	1 MKGATLVALAATIGNFLOGWDNATIAGAMVYINKDNL--PTSVQGLVAMSLIGATVI	57
QY	115 TTFSGPLSDSIGRRPMLILSSILYFFSGLIMLMSPNVYVLLARFVDGFGIGLAATVPL	174
Db	58 TTCSGPISDWLGRPMLILSSVMYFVCGGLIMLMSPNVYVLCFARLLNGFAGLAATLVPL	117
QY	175 YISEIAPSEIRGLNTLPQFSGSGGMFLSYCMVFEGMSLSPSPDWIRIMLGVALIPSLFFFG	234
Db	118 YISETAPPEIRGQNLTLPPQFLSGGMFLSYCMVFTMSLSDSPSWRAMLGVALIPSLLYLF	177
QY	235 LTFYLPESPRWLVSKGMAEAKVLOKLRKDDVSGELSLLEGLVEGDTISIEYIIG	294
Db	178 LTFYLPESPRWLVSKGRMDEAKRVLOQLCGREDVTGMALLVEGLDIGEKTMDLLVT	237
QY	295 PATEAADLVTDGKE-QITLYGPEGQSWIARPSKGPIMLGSVLSLARHGSMTVNSVP	353
Db	238 LEDHGGDTLETVDEGQMRLYGTHENQSYLARVPEQ--NSSLGRSRHGSLSANQSMI	294
QY	354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTQ--HAKNEQWD--EENLHRD	408
Db	295 LKDPVLNLFGSLHEKMPAEGCNTRSGIFPHFGSMFSTTADAPHGKPAHWEKDIESHYNKD	354
QY	409 DEEYASDGAGDYED--NLHSPILSRQATGAEGKDIVHGHRSALSRRQTLLEGGCD	465
Db	355 NDDYATDDGAGDDDDSDNDLRSPLMRSRQTTSM-DKMIIPHPTSGSTLSMRHSTLMQ-GN	412
QY	466 GVSSTDIGCGWQLAWKMKSEKEGNGRKEGKRVYLHQEGVPGSRGSIIVSLPGGDFVE	525
Db	413 GESSMGIGCGWHMGYRYENDE-----YKRYLLKEDGAB-SRGSITISIPGPD--G	460
QY	526 GSEFVHAALVQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGV	585
Db	461 GGSYIHASALVRSVLGPKS----VHGSAMVPEKIIAASGPLWSALFEPGVRRALLVGV	515
QY	586 GIQILQGFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTTLLMLPCIGFA	645
Db	516 GIQILQFGSINGVLYTTPQILERAGVDILSSIGLSSISASFLISGLTLLMLPAIVVA	575
QY	646 MLMDLSGRRFLLGTIPILIASLVILVNSNLDLGLAHALLSTVSIVYFCCFVMGFG	705
Db	576 MRLMDVSGRRSLTLMTIPVLIVSLVVISLIIHISKVVNAALSTGCVVLYFCFFVMGYG	635
QY	706 PIPNIIICAEIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLI	765
Db	636 PIPNIIICSEIFPTRVRGLCIAICAMVFWIGDIIIVTYSLPVLSSIGLVGVFSIYAAVCVI	695
QY	766 SFVFLVFLKVPETKGMPLLEVITEFFAVGAKQA 800	

Db 696 SWIFVYMKVPETKMPLEVITDYFAFGA-QAQASA 729

RESULT 6

Q9LM67 PRELIMINARY; PRT; 734 AA.
AC Q9LM67;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE F2D10.36.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bel O., Chin C., Chio J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Tortum M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F2D10 from chromosome
RT I.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AC069251; AAF80627.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 734 AA; 79485 MW; 0232FE8BF850259F CRC64;

Query Match 55.8%; Score 2275; DB 10; Length 734;
Best Local Similarity 61.2%; Pred. No. 3.1e-146;
Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;

QY 56 MSGAVLVAIVASIGNLLQGMNATIAAVALYIKKEFOQNEPT-VEGLIVMSLIGATIV 114
DB 1 MKGATLVALAATIGNFLQGMNATIAAMVYINKDNL-PTSVOGLVAMSLIGATIVI 57
QY 115 TTFSGPLSDSIGRRPMLISSILYFSGLIMLWSPNVYLLARFVDGFGIGLAVTLVPL 174
DB 58 TTCSGPISDWLGRRPMLISSVMYFVCGLIMLWSPNVYVLCFARLLNGFGAGLAVTLVPV 117
QY 175 YISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPSLFFFG 234
DB 118 YISETAPPEIRGQLNLTLPGLSGGMFLSYCMVFTMSLSDSPSRAMLGVLSIPSLYLF 177
QY 235 LTIFYLPESPRMLVSKGRMAAKVYLQKRGKDDVSGELSLLEGLVEGGDTISEYIIG 294
DB 178 LTVFYLPESPRMLVSKGRMAAKVYLQKRGKDDVSGELSLLEGLVEGGDTISEYIIG 294
QY 295 PATEAADLVTDGKE-QITLYGPEEGQSWIARPSKPIMLGSLVSLASRHGSMTVQSV 353
DB 238 LEDHEGDDTLETVDDEGQMLRYGTHENQSYLARVPPEQ--NSSGLRSRHGSLANQSMI 294
QY 354 LMDPIVTLFGSVHENMPQAGSMESTLFPNFGSMFSVTDQ--HAKNEQWD--EENIHRD 408
DB 295 LKDPVLNLFGLHEKMEPEAGNTRSGIFPHFGSMFSTADAPHGKPAHWEKDIESHYKND 354
QY 409 DEEYASDAGAGDYED--NLHSPLLSRQATGAEGKDIYHGHRSALSMRRQTLLEGCD 465

Db 355 NDDYATDDGAGDDDDSDNDLRSPLMSRQTTSM-DKMIPHPTSGSTLSMRHSTLMQ-GN 412
QY 466 GVSSTDIGGQWLAKWKSEKEGENGRKEGGRVYLHOEGVPGSRGSIYSLPGGDVFE 525
DB 413 GESSMGIGGGMHGYRYENDE-----YKRYLKEDAE-SRGRSITISIPGPD--G 460
QY 526 GSEFVHAALVSGALFSKGLAEPKSDAAMVHPSEVAAKSGRWKDLFEPGVRALLVGV 585
DB 461 GGSYIHASALVSRVLPKPS----VHGSAWVPEKTAASGPLWSALBEPVKRALVGV 515
QY 586 GIQILQOFAGINGVLYYTPQILEAGVAVISKFGSSASASILISLTLLMLPCIGFA 645
DB 516 GIQILQFSGINGVLYYTPQILERAGVDILSSGLSSISASFLISGLTLLMLPAIIVA 575
QY 646 MLMDLSGRFLLGTIPILIASLVILVSNLIDGLTALHALLSTVSVIYFCCFWMGFG 705
DB 576 MRLMDVSGRSLMLTIPVLIVSLVVLVISELIHISKVNAALSTGCVLVYCFVMGYG 635
QY 706 PIPNLCAEIPFTRVRLCIAICAFTEWIGDIIVTYSLPVMLNAGLACVSIYAVCLI 765
DB 636 PIPNLCSEIFPTRVRLCIAICAMFWIGDIIVTYSLPVLLSSIGLVGFSIYAAVCVI 695
QY 766 SFVFPFLKVPETKMPLEVITEFFAVGAQAQAAKA 800
DB 696 SWIFVYMKVPETKMPLEVITDYFAFGA-QAQASA 729

RESULT 7

Q9SD00 PRELIMINARY; PRT; 729 AA.
ID Q9SD00
AC Q9SD00;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Sugar transporter-like protein (Monosaccharide sensing protein
DE 3).
GN F26013.130 OR MSSP3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Delseney M., Berger C., Cooke R., Grellet F., laudie M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Stamme C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
RT "A novel family of plant monosaccharide transporters is involved in a
RT new type of eukaryotic sugar sensing.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AL133452; CAB63013.1; -.
DR EMBL; AJ532571; CAD58693.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 729 AA; 78828 MW; 241A9415DBA343FCCC CRC64;

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eucosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Tjaden J., Neuhaus E.;
RT "A new sugar transport protein from Arabidopsis thaliana.";
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; Z50752; CAA90628.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 734 AA; 79462 MW; 1B009F5F457AF6E1 CRC64;

Query Match	53.6%;	Score 2183;	DB 10;	Length 734;
Best Local Similarity	59.5%;	Pred. No. 5.7e-140;		
Matches 449;	Conservative 110;	Mismatches 160;	Indels 36;	Gaps 14;

```

QY      56 MSGAVLVAIVASIGNLQQWMDNATIAAAVLVIKKEFQLQNEPT-VEGLIVSMSLGATIV 114
      | || ||| : ||||| ||||| ||| : ||| : || ||||| ||||| : ||||| :
Db      1 MKGATLVALAATIGNFLQGWMDNATIAAGAMVYINKDNL--PTSVQGLVVAWSLIGATIV 57

```

```
QY      115 TTFSGPLSDSIGRRPMLILSSILYFSGILMLMSPNVVLLLARFVDGFGIGLAVTLVPL 174
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      58  TTCSGPISDWLGRRRPMLILSSVMYFVCGILMLMSPNVVLCFARLLNGFGAGLAVTLVPV 117
```

```
Qy      175 YISEIAPSEIRGLNTLTPQFSGSGGMFLSYCMVFQMSLSPSPDWRIMLGVTALPSLFFFC 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db      118 YISETAPPEIRGQNTLTPQFLSGGGMFLSYCMVFQMSLSDSPSWRAMLGVTSLPSLLYLF 177
```

Qy	235	LTFYLLPESPRWLVS	SKGRMAEAKKVT	OKLRKQDVSGEL	SLLEGLVGGDTS	IEEYII	G	294
		:		:	:	:	:	
Db	178	LTFYLLPESPRWLVS	SKGRMDEAKRVL	QOLCGREDVT	DEMALLVEGLD	IGGEKTMED	LVT	237

```
Qy      295 PATEAADDLVTGDKE-QITLYGPEEGSQWIARPSKGPIMLGSVLSLARHGSWVNOQVP   353
        || : ||| | |||::||| | | | ||||| : ||| :
Db      238 LEDHEGDDTLFETVEDEGGQIRLYGTHENOSYIARPVPEQ--NSSLGLRSRHGSILANOSMI   294
```

```

QY      354  LMDPIVTLFGSVHENMPQAGGSMRSTLPNFGSMFSVTDQ--HAKNEQWD--EENLHRD 408
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      295  LKDPVLNLFGLHEKMPGAGNTRSGIFPHFGSMFSTTADAPHGKPAHWKOIDESHYNKD 354

```

QY 409 DEEYASDAGGDYED---NLHSPLLSRQATGAEGKDIVHGHGSGALSMRRQTLLEGGD 4655

Db 355 NDYATDDGAGDDDDSDNDLRSPLMSRQTTSM-DKMIPHPTSGTSLSMRRHSTLMQ-GN 4122

```

Qy      466 GVSSTDIGGGWQLAWKWSKEGENGKKEGFKRYVLHQEGVPGSRRGSIIVSLPGGGVFE 525
          ||| |||| : :: | : || : || |||| : |||| |
Db      413 GESSMIGGGWMMGYRYENDE-----YKRYLLKEDGAE-SRRGSIISIPGGPD--G 460

```

```
QY      526 GSEFHAALVSGSALFSKGLAEPMSDAAMHSEVAAKGSRWKDLFEPCVRALLVG 585
        |::|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      461 GGSIHASALVSRSVLGPKS-----VHGSAWPEPKETAAAGPLWSALLEPGVKRALVGV 515
```

```

QY      586  GIQLIQFAGINGVLYTTPQILEAGVAVILSKFGSSASASILISSLTTLMLPCIGFA 645
          |||||:|||||||:||||:|||||
DB      516  GIQLIQFSGINGVLYTTPQILERAGVDILSSLGSSISASFLISGLTTLMLPAIYVA 575

```

Oy 646 MLMLDSCGRFLLLTIPILIASLVTLVNSNLDCTLAHALSTSVIVYFCFPWGERG 705
| | | : | | | : | | | : | | | : | | | : | | | : |
Db 576 MRLMDVSGRRSLLMTIPLVLISLVTVLISELIHISKVNAAALSTGCVVLYFCFPWGYG 635

QY 706 PIPNLCAEIPTRVRCGLCIAICAFTFWIGDIIVTYSLPMLNAGLAGVESIYAVVCLT 7655

Db PFQTSVLKSSQQADRGCLCIAICAVFWIGDIIVTYSLPVLSSIELVGFSIYAAVCVI 695

QY SFVFVFLKVPETKGMPLLEVITTEFFAVGAKQQAAGA 800

Db SWIFVYMKVPETKGMPLLEVITDYFAFGA-QAQAASA 729

RESULT 10
Q8H887
ID Q8H8

PRELIMINARY; PRT; 714 AA.

DT 01-MAR-2003 (TREMBlere1. 23, Created)

DT	01-MAR-2003 (TREMBlé, 23, Last annotation update)
D1	01-MAR-2003 (TREMBlé, 23, Last sequence update)

DE putative sugar transporter protein.
GN OJ1528D07.4.

OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Uryza sativa (japonica cultivar-group) .

0C Ebrhartioideae: Orvzeae: Orvzā.
0C Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OX	NCBI_TaxID=39947;
RN	[1]

RP SEQUENCE FROM N.A.
PA Wind P A Y1 Y

Currie J., Collura K.;
"Dina Canavia Canavia"

Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

KW	Sugar transport.
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99

Query Match	43.3%	Score 1763;	DB 10;	Length 714;
Best Local Similarity	47.5%	Pred. No. 2.1e-111;		
Matches 377; Conservative	125;	Mismatches 193;	Indels 98;	Gaps 14;

OY 4 GSWLAVQTPTTPDLLRRERLLPSVVLALFEPGLPPASCSSQEPVTSDLDLEDKMSGAVLVA 63
||::|::|::|::|
9 GAW-----KNTROQYSPSPSPLSGSSVFSESSLSTTAD---GSTANAVLAA 54

Db

QY	64	I V A S I G N L L Q G M D N A T I A A A V L Y I K K E F Q L O N E P T V E G L I V S M S L I G A T I V T T F S G P I S D 123
		: : : : : : : :
Db	55	I V A S I G N L L Q G M D N A I A G A I M Y I K N E F N L O N D P M E E G L I L A M S L I G A T I T I T A L S C M I T N 114

```

QY      124 SIGRRPMLILSSILYFFSGLIMLWSPNYVLLARFVDGFIGLAVTLVPLYISEIAPSE 183
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::
Db      115 SIGRRPILLSVAALILYSISALIMFQASNEFMILLARLIYGFSGSLVVTYAPLYISEIAPTN 174

```

```

Qy      184  IRGLINTLPQFSGSGGMFLSYCMVFEGNSLSPSPDWRIMLGVLAIPLFFFGTLTFYPES 243
      :|||||||:|||||||:|||||||:|||||||:|||||||
Db      175  MRGLINTLPQFNGSLGMLLSYIMVFLMSLTLPNPNWRIMLGSLSPSFVFLLCIFYLPES 234

```

QY	244	PRWLVS	SGRMAE	AKVLQ	LRGKDP	VS	GELSLL	LEGL	EV	GGDT	SI	EYII	IG	PATEA	ADDL	303
			:			:		:		:		:		:		
Db	235	PVFLV <th>SGKIE</th> <th>BAKVM</th> <th>KRLRG</th> <th>TNEVS</th> <th>SEIAFL</th> <th>IGLTV</th> <th>DQDNY</th> <th>IEDY</th> <th>MGHN</th> <th>DEED</th> <th>-</th> <th>293</th>	SGKIE	BAKVM	KRLRG	TNEVS	SEIAFL	IGLTV	DQDNY	IEDY	MGHN	DEED	-	293		

```

Qy      304 VTDDGKEQITLYGPEEGQSWIARPSKPIMLGSLSLASRHGSMVNQSVPLMDPIVTLFG 363
      : : | | | | : | | | : | : | | | | |
Db      294 QSISNETETTKLYGHEEGVTWEARPFKKNV-----VESDHSPIPN---LLDPIVTLFD 343

```

Oy		364	SVHENMPEAGGSMRSTLFPNFGSMPFVTDQAKNEQWDEENLRHDDDEEYASDGAGGYED	423
			: : : : : :	
Dd		344	SIHGN-----LNTPEFTSSGNMSNDIEQ-PTKDLDSQEDLDLT-----DYED	384

OY 424 NIHSPLLRSQATGAEGKDIVHGHGRGSALSMRRQTLLGEGGDVSTDIGGWOLAWKWS 483

DB 385 DLGHPLLF-----HGGSYM-----EGIDDACVNGGWHIAWKFEV 417

```

QY      484  EKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGCVFEGSEFVHAALVSQSALES 543
      ::|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      418  QRENFEGQTDDFQQIFLQ-----GDILQAGRVSHATALVSTPS--FH 456

```

QY 544 KGLAEPRMSDAMVHPSE--VAAKGRNQLFEPGVRALLVGVGIIQLQGFAGINGVLY 601


```
Db 459 HSLG-----PAMVHPSKFNLSSTEGQSWSDLQPGVKQGLIVGTIQLIQLAGISGILY 512
QY 602 YTPQILEQAGAVAILSKFGLSSASASILISLTLLMPCIGFAMLLMDLSGRFLLIGT 661
Db 513 YTFQILEQAGAGILKWFENVSSSSSILTSALTFTMLPSIGIAMKCMDRYGRSLLYT 572
QY 662 IPILIASLVILVSNLIDGLTAHALISTVSIVVFCCFVNGFGPIPNILCAEIPTRVR 721
Db 573 IPMLIVSLIILIVNVNMLEAIFGAILSTFGVIIVCCFVNGFGPIPNVLCSELPSCSR 632
QY 722 GLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSYAVVCLISFVFLKVPETKGM 781
Db 633 NRCMSICTLTFWIVSIIVTYAFPVMLSSIGLIGVCGIYAVVCIVSFIFVLKVPETKGM 692
QY 782 LEVITEFFAVGAK 794
Db 693 LAVIANSLAVGAR 705

RESULT 11
Q9SNK7 PRELIMINARY; PRT; 723 AA.
ID Q9SNK7
AC Q9SNK7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to sugar transporter protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(Ga3) genomic DNA, chromosome 3, PAC
RT clone:P0043E01.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AP000615; BAA85398.1; -.
DR Gramene; Q9SNK7; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 723 AA; 79061 MW; EC1E924FC4206BD8 CRC64;

Query Match 43.3%; Score 1763; DB 10; Length 723;
Best Local Similarity 47.5%; Pred. No. 2.1e-111;
Matches 377; Conservative 125; Mismatches 193; Indels 98; Gaps 14;

QY 4 GSWLAVQTPFTPDLDREERLLPSVLLALPGPLPASCSSQEPVTSDDLLEDKMSGAVLVA 63
Db 18 GAW-----KNTROQYSSPSSPISLSSSVFSESSLELSTTTAD----GSTANAVLAA 63
QY 64 IVASIGNLQGWDMNATIAAVALYIKKEFQLONEPTVEGLIVSMSLIGATIVTTFSGPLSD 123
Db 64 IVASIGNLQGWDMNATIAAVALYIKNEFNLQNDPMMEGLILAMSLIGATITITALSGMTN 123
QY 124 SIGRRPMLILSSILYFSGLIMLWSPNVVLLARFVDFGIGLAVTLVPLYSEIAPSE 183
Db 124 SIGRRPMLSVAAILYISALIMFQASNEYMLLRLIYFGSGLVVTYAPLYSETAPT 183
QY 184 IRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFFGTLTFYLPES 243
```

```
Db 184 MRGLNLTLPQFNGSLGNLSTYIMVFLMSLTLPNWRIMLGLSLISPSFVFLLCIFYLPES 243
QY 244 PRMLVSKGRMAEAKYLOKLRKDDVSGELSLLEGLVVGCDTSEIYYITGPATEAADL 303
Db 244 PVFLVSKKIEEAKVNMKRLRGTEVSSSEIAPLIQGLITVDQNDYIEDYMGHNDEPDD- 302
QY 304 VTDGKEQITLYGPEQSWIARPSKPIMLGSVLSARHGSVMNQSVPLMDPIVTLFG 363
Db 303 QSISNTETTKLYGHEBGVTWFAFPFKGNV-----VESDHSPIPN---LDPDIVTLFD 352
QY 364 SVHENMPQAGSMRSTLPNFGSMFSVTDOHAKNEQWDEENLHRDEEYASDAGGDYED 423
Db 353 SIHGNI-----LNTPEFTSSGNMSNDIEQ-PKTDLESQEDLDT-----DYED 393
QY 424 NLHSPILSRQATGAEKDIVHGHGSAALSMRRQTLLEGGDVSTDIQGGWQLAWKWS 483
Db 394 DLGHPLLF-----HQSVM-----EGIDDACVNGWHIAWKV 426
QY 484 EKEGENRKEGFRKRYLHQEGVPGSRGSIYSLPGGDFEGSEFVHAALVSQSALFS 543
Db 427 QRENEFGQTQDDFQOIFLQ-----GDILQAGRVSHATALVSTPS-FH 467
QY 544 KGLAEPRMSDAAVHPSE--VAKGSRWKDLFEPGVRALLVGVGIOILOQFAGINGVLY 601
Db 468 HSLG-----PAMVHPSKFNLSSTEGQSWSDLQPGVKQGLIVGTIQLIQLAGISGILY 521
QY 602 YTPQILEQAGAVAILSKFGLSSASASILISLTLLMPCIGFAMLLMDLSGRFLLIGT 661
Db 522 YTPQILEQAGAGILKWFENVSSSSSILTSALTFTMLPSIGIAMKCMDRYGRSLLYT 581
QY 662 IPILIASLVILVSNLIDGLTAHALISTVSIVVFCCFVNGFGPIPNILCAEIPTRVR 721
Db 582 IPMLIVSLIILIVNVNMLEAIFGAILSTFGVIIVCCFVNGFGPIPNVLCSELPSCSR 641
QY 722 GLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLKVPETKGM 781
Db 642 NRCMSICTLTFWIVSIIVTYAFPVMLSSIGLIGVCGIYAVVCIVSFIFVLKVPETKGM 701
QY 782 LEVITEFFAVGAK 794
Db 702 LAVIANSLAVGAR 714

RESULT 12
Q8LJX7 PRELIMINARY; PRT; 746 AA.
ID Q8LJX7
AC Q8LJX7;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Putative sugar transporter.
GN SB234M12.4.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoidae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. BTx623;
RA Song R., Liaca V., Messing J.;
RT "Mosaic Organization of Orthologous Sequences in Grass Genomes.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF527809; AAM94321.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 2.
DR PRINTS; PR00171; SUGRTNSPORT.
```


RESULT 14

Q9XEV1 PRELIMINARY; PRT; 664 AA.
AC Q9XEV1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Putative sugar transporter protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Llaca V., Lou A., Young S., Messing J.;
RT "Microsynteny in a gene-dense region in maize, sorghum, and rice."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF119222; AAD27676.1; -.
DR Gramene; Q9XEV1; -.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Sugar transport; Transmembrane.
SQ SEQUENCE 664 AA; 71851 MW; BA0CAE40C11C29E7 CRC64;

Query Match 15.3%; Score 621.5; DH 10; Length 664;
Best Local Similarity 27.5%; Pred. No. 8.9e-34;
Matches 193; Conservative 87; Mismatches 183; Indels 239; Gaps 21;

QY 1 IRSGSWLAVQTPFT-----PDLDRERLLPSVIALPGLPPASCS----- 41
Db 123 VRTEPOLCSSTPSTNHQSSANQQLPSPOSHRLRLHPRAI-----PLPPAIAARSKLASP 178
QY 42 -----SOEPTVSD-----ILEDKMSGAVLVAIV 65
Db 179 RYRGTGAPSFRRGAGARIIPSEVAPSDRPNHGESSNYLERAIVLDEMKSSTVSAVA 238
QY 66 ASIGNLLQGWDNATI-----AAVLYIKKEFOLQNEPTEVGLIVMSLIGATVTPSGP 120
Db 239 VSIGYTLGWDFTVLGNPSAEANIHMKKEFGLNCGPSIDGILLAVSVGSIAITVSGS 298
QY 121 LSDSIGRRPMLILSSILYFSGILMLMSPNVVLLARFVDFGIGLAVTLVPLTSEIA 180
Db 299 LLDMLGRR----- 306
QY 181 PSEIRGLNTLPQSGSGMFLSYCMVFGMSLSPSFDWRIMLGVLAIPLSFFGLTIFYL 240
Db 307 -----IVFSYCLIFWMTLIPSFNWRIMGAIFAPSLVYFALIVYL 347
QY 241 PESPRWLVSKGMAEAKKVLQKRGKDDVS-----GELSLLEGLEVGDTSIEEYITGPA 296
Db 348 PESPRWLVSQKISEARISLQWLRGKDDVSDRSAGEIALIADGMNMITETAVGHAIVGAV 407
QY 297 TEAADLVTGDKKEQITLYGPERGQSWIARPSKGPIMLGSVLSLASRHSVMNQSVPLMD 356
Db 408 -----RSQSF-----LGTSTNQMSRSHSTFYWH---LSD 432
QY 357 PIVTLFGSVHNMPOAGSMRSTLFPNFGSMFSVTDOHAKNEQWDEENLRHDEEYASDG 416
Db 433 PLVDLLGSIHESMSELAG-RNSYFVFNNS-FNIVEQEQTSEQRGNDLSIQSREAYSAE- 489
QY 417 AGGDYEDNLHSPILSRQATGAEGKI-VHGHGSGALSMPROTLLGEGGDVSTDIGG 475
Db 490 -EGNNGDNLQASLLS-QVASAETNDINTSFTSESSSSYLRH---GTSTSGLAQDLISS- 543
QY 476 WQIAWKSEKEGNGRKEGFRKRYVLHQEGVPGSRGSIIVLPGGGVFEGSEFYHAAL 535
Db 544 -----LHDH-----DIEEDEEIHIAL 561
QY 536 VSQSALFSKGLAEPKMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIIQLQFAG 595

Db 562 SSQPAL-GAGLHPRQO---MRLSETADIKPKWRVLLQPGVRHALCYGMILQALQO--- 614
QY 596 .INGVLYTTPQILEQAVAVILSKFGLSSASASILISLTLL 637
Db 615 -----TFNLLPGDAKYHWAQWGMCDLCIRVLLGSRVRL 650

RESULT 15

P96742 PRELIMINARY; PRT; 457 AA.
AC P96742;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to metabolite transport proteins.
GN YWTG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Lelong C., Glaser P., Prescecian E., Danchin A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,
RA Brouillet S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Grandi G.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Gutseppl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones J.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwollik S., Prescott A.M.,
RA Prescecian E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis";
RL Nature 390:249-256 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; Z92954; CAB07473.1; -.
DR EMBL; Z99122; CAB15600.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.

DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Complete proteome.
SQ SEQUENCE 457 AA; 49192 MW; E688ECE04C8B6EF7 CRC64;

Query Match 14.4%; Score 587.5; DB 16; Length 457;
Best Local Similarity 22.8%; Pred. No. 1.1e-31;
Matches 168; Conservative 102; Mismatches 162; Indels 305; Gaps 12;

QY	55	KMSGAVLAIVASIGNLQGDNATIAAAVLYIKKEFQLQNEPTVEGLIVMSLIGATIV	114
DB	2	KKQSNIMLYFPGALGALGYDYGVISGAILFMKKEGL--NAFTEGLVSSLLVGAILG	59
QY	115	TTFSGPLSDSIGRRPMLLISSILYFSGLIMLMSPNVYVLLARFVDGFGIGLAVTLVPL	174
DB	60	SGAAGKLTDRFGRKKAIMAAALLFCIGGLVALAPNTGVMLFRIILGLAVGTSTTIVPL	119
QY	175	YISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVEGMSLSPSPDWRIMLGVALPSPFPFG	234
DB	120	YLSELAPKHKRGALSSLNQIMITVILSY--IVNYIFADAEMRWMLGLAAVPSLLLL-	176
QY	235	LTIFYLPESPRWLVSKGMAEAKVLQKLRKQDVSGELSLLEGLVGGDTSEIYYIG	294
DB	177	IGLFMPESPRLFTNGESKAKILEKLRGTQDI-----	211
QY	295	PATEAADLVTGDKQITLYGPEEGQSWIARPSKPIMLGSLASRHGSMVNGSVPL	354
DB	212	-----	211
QY	355	MDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDDEEYAS	414
DB	212	-----DQE---	214
QY	415	DGAGGDYEDNLHSPLLSRQATGAEGKDIVHGHGSGALSMRQOTLGEGGDVSSTDIGG	474
DB	215	-----IH-----DI--	218
QY	475	GWQLAWKWEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVLPGGDVFEGSEFVHAAA	534
DB	219	-----KEAEKQDEGGKE-----	231
QY	535	LVSQSALFSKGLAEPRMSDAAVHPSEVAAKGSRWKDLFEPGVRALLVGVGIQLIQFA	594
DB	232	-----LFDPMWRPALIAGLGLAFLQQFI	254
QY	595	GINGVLYYTPQILEQAGVAVILSKFGLSSASIL-----ISSLTLLMLPCIGFAMLLMD	650
DB	255	GTNTIIYYAPKFTNVG-----FG--NSASILGTVGIGTVNVMTL---VAIKIID	300
QY	651	LSGRKFLLLGTIPILIASLVILVSNLIDGLTAHALSTVSIVIVYFCCFVMGFGPIPNI	710
DB	301	KIGRKPLLLFGNAGMVISLIVLALVNLFPNNTPAASWTIVICLGVPIVFAVSWGPPVWV	360
QY	711	LCAEIPPTRVRGLCIACAFTEWIGDIIVTYSLPVMLNAGLAVFSIYAVVCLISFVFV	770
DB	361	MLPELPLHVRGIGTGVSTLMHVTLVSTYPIIMEAIGISYFLIYAAGIMAFLEFV	420
QY	771	FLKVPETKGMPLLEVITE 787	
DB	421	RFKVTETKGRSLBEIEQ 437	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 15:48:00 ; Search time 74 Seconds
(without alignments)
1715.963 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPDLDRL.....PLEVITTEFFAVGAKQAATAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 19Jun03:*

1:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA2001.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA2002.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneeqp-emb1/AA2003.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3517	86.3	747	23	AAU97201	Corn Arabidopsis t
2	3517	86.3	747	24	ABU08326	Corn sugar transpo
3	2674	65.6	737	23	AAU97204	Soybean Arabidopsi
4	2674	65.6	737	24	ABU08329	Soybean sugar tran
5	2584	63.4	729	23	ABB93252	Herbicidally activ
6	2275	55.8	734	23	ABB91078	Herbicidally activ
7	2245	55.1	729	23	ABB92622	Herbicidally activ
8	1670.5	41.0	486	23	AAU97205	Portion of a soybe
9	1670.5	41.0	486	24	ABU08330	Soybean sugar tran

10	1424	34.9	345	23	AAU97207
11	1424	34.9	345	24	ABU08332
12	943	23.1	228	23	AAU97208
13	943	23.1	228	24	ABU08333
14	562	13.8	131	23	AAU97203
15	562	13.8	131	24	ABU08328
16	554.5	13.6	487	24	ABG73334
17	548	13.4	488	24	ABP98504
18	520.5	12.8	555	21	AAG29529
19	520.5	12.8	582	21	AAG29528
20	496	12.2	523	23	AAU97212
21	496	12.2	523	24	ABU08337
22	493.5	12.1	491	21	AAG32072
23	493.5	12.1	508	21	AAG32071
24	489.5	12.0	513	23	AAU97210
25	489.5	12.0	513	24	ABU08335
26	489	12.0	529	23	AAU97214
27	489	12.0	529	24	ABU08339
28	485	11.9	510	23	AAU97211
29	485	11.9	510	24	ABU08336
30	481	11.8	513	23	ABB09681
31	468	11.5	551	20	AAV50799
32	467	11.5	478	21	AAG15416
33	467	11.5	493	21	AAG15415
34	467	11.5	546	21	AAG15414
35	463.5	11.4	539	23	AAU97213
36	463.5	11.4	539	24	ABU08338
37	457.5	11.2	466	21	AAG32073
38	453.5	11.1	480	22	AAM00107
39	451	11.1	464	23	ABP52163
40	448	11.0	474	19	AAW58862
41	444	10.9	491	22	AAG89949
42	443	10.9	508	21	AAG29530
43	442	10.8	491	21	AAB12594
44	429.5	10.5	472	23	ABP52164
45	425	10.4	629	22	AAB66935

ALIGNMENTS

RESULT 1	
AAU97201	
ID	AAU97201 standard; Protein; 747 AA.
XX	
AC	AAU97201;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Corn Arabidopsis thaliana-like sugar transport protein.
XX	
KW	Corn; Arabidopsis thaliana-like sugar transport protein;
KW	carbohydrate transport; grain filling; annual field crop;
KW	plant.
XX	
OS	Zea mays.
XX	
FH	Key
FT	Misc-difference 129
FT	Misc-difference 133..134
FT	Misc-difference 133..134
FT	Misc-difference 144
FT	Misc-difference 144
FT	Misc-difference 178
FT	Misc-difference 178
FT	Misc-difference 207
FT	Misc-difference 207
FT	Misc-difference 218
FT	Misc-difference 218
FT	Misc-difference 220
FT	Misc-difference 220
FT	Misc-difference 236
FT	Misc-difference 236

Portion of a wheat
Wheat sugar transp
Portion of a wheat
Wheat sugar transp
Portion of a rice
Rice sugar transpo
Consensus amino ac
PFAM consensus seq
Arabidopsis thalia
Arabidopsis thalia
Soybean Beta vulga
Soybean sugar tran
Arabidopsis thalia
Arabidopsis thalia
Corn Beta vulgaris
Corn sugar transpo
Wheat sugar transp
Wheat sugar transp
Rice Beta vulgaris
Rice sugar transpo
Amino acid sequenc
Spinach glucose tr
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Wheat sugar transp
Wheat sugar transp
Arabidopsis thalia
Sugar transporter
E. coli galactose-
T. halophilus xyul
C glutamicum prote
Arabidopsis thalia
Brevibacterium lac
E. coli arabinose-
Human GLUTX2. Hom

Application

FT	/label= unknown	
XX		
PN	US6383776-B1.	
XX		
PD	07-MAY-2002.	
XX		
PF	14-APR-1999;	99US-02919922.
XX		
PR	24-APR-1998;	98US-083044P.
XX		
PA	(DUPO) DU FONT DE NEMOURS & CO E I.	
XX		
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;	
XX		
DR	WPI; 2002-451386/48.	
DR	N-PSDB; ABK51962.	
XX		
PT	New nucleic acid encoding plant sugar-transport proteins, useful for preparing transgenic plants with altered carbohydrate distribution -	
XX		
PS	Claim 1; Fig 1; 54pp; English.	
XX		
CC	The present invention relates to the isolation of plant polynucleotide sequences encoding an Arabidopsis thaliana-like sugar transport protein or Beta vulgaris-like sugar transport protein. The polynucleotide sequences are useful for altering the level of sugar transport proteins in plants, i.e. for control of carbohydrate transport and distribution in plant cells, e.g. during grain filling of annual field crops (e.g. corn, rice, soybeans, and wheat), and, for studying carbohydrate flows and sugar transport. The polynucleotide sequences can also be used to isolate cDNA sequences and genes that encode homologues of the new proteins. The present sequence represents a corn Arabidopsis thaliana-like sugar transport protein.	
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		

Db	481 ALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIIQLQOF 540
QY	594 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 653
Db	541 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 600
QY	654 RRELLGTIPILIASLVILVSNLIDLTAAHALLSTVSVIVYFCCFVWGFGPIPNILCA 713
Db	601 RRELLGTIPILIASLVILVSNLIDLTAAHALLSTVSVIVYFCCFVWGFGPIPNILCA 660
QY	714 EIEPTRVRGLCIACAFTFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLK 773
Db	661 EIEPTRVRGLCIACAFTFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLK 720
QY	774 VPETKMPLEVTFFFAVGAKQAAAKA 800
Db	721 VPETKMPLEVTFFFAVGAKQAAAKA 747

Application

RESULT 2	
ABU08326	
ID	ABU08326 standard; Protein; 747 AA.
XX	
AC	ABU08326;
XX	
DT	29-MAY-2003 (first entry)
XX	
DE	Corn sugar transport protein #1.
XX	
KW	Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW	plant sugar transport protein; carbohydrate transport; soybean;
KW	carbohydrate distribution; plant.
XX	
OS	Zea mays.
XX	
PN	US2002178468-A1.
XX	
PD	28-NOV-2002.
XX	
PF	17-JAN-2002; 2002US-0051902.
XX	
PR	24-APR-1998; 98US-083044P.
XX	
PR	14-APR-1999; 99US-02919922.
XX	
PA	(ALLE/) ALLEN S M.
PA	(HITZ/) HITZ W D.
PA	(KINN/) KINNEY A J.
XX	(TING/) TINGEY S V.
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX	
DR	WPI; 2003-340957/32.
DR	N-PSDB; ABX93198.
XX	
PT	Novel plant sugar transport proteins and nucleic acid encoding the
PT	protein useful for producing transgenic plants having altered levels of
PT	sugar transport protein -
XX	
PS	Claim 5; Fig 1; 56pp; English.
XX	
CC	The present invention relates to the isolation of Arabidopsis
CC	thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC	the polynucleotide sequences encoding them. The plant sugar transport
CC	proteins of the invention have been isolated from corn, rice, soybean,
CC	and wheat. The polypeptides of the invention may be used for altering
CC	the level of expression of a sugar transport protein in a host cell,
CC	by transforming a host cell with a chimeric construct encoding all,
CC	or a portion of the sugar transport protein, in sense or antisense
CC	orientation. Particularly, the polypeptides may provide a means to
CC	control carbohydrate transport and distribution in plants.
CC	ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
CC	transport proteins.

XX Sequence 747 AA;
SQ Query Match 86.3%; Score 3517; DB 24; Length 747;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY	56	MSGAVLVAIVASIGNLQGMWDNATIAAVALYIKKEFQONEPTVEGLIVSMGLIGATIVT	115
Db	1	MGAVMVAIAASIGNLQGMWDNATIGAVLYIKKEFNIOSEPLIEGLIVAMFLIGATIVT	60
QY	116	TFSGPLSDSIGRRPMLILSSILYFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY	175
Db	61	TSPGPRADCVGRPRMVLVASAVLVFVSGLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY	120
QY	176	ISEIAPSEIR-GLNLTLFQFSG-SGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLFF	233
Db	121	ISETAPHRXSWGXXNTLPQFIGVXGGMFLSYCMVFGMSLMPKPDWRIMLGVLAIPLFF	180
QY	234	GLTIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIYYII	293
Db	181	GLTVFYLPESPRWLVSKGMAEAKVXQRLRGREDVSEXALLVEGLVGKDTIRIXEYII	240
QY	294	GPATEAADLVTDGKEQITLYGPEEGQSWIARPSKPIMLGSLVSLASRHGSMVNOQSV	353
Db	241	GPATEAADLVTDGKEQITLYGPEEGQSWIARPSKPIMLGSLVSLASRHGSMVNOQSV	300
QY	354	LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEYA	413
Db	301	LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEYA	360
QY	414	SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRGSAISMROTLIGEGGDGVSSTDIG	473
Db	361	SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRGSAISMROTLIGEGGDGVSSTDIG	420
QY	474	GGWOLAMKWEKEGNGRKEGFKRVYLHOGVPGSRGSIIVSLPGGGDVFESEFVHAA	533
Db	421	GGWOLAMKWEKEGNGRKEGFKRVYLHOGVPGSRGSIIVSLPGGGDVFESEFVHAA	480
QY	534	ALVSQSALFSKGLAEPKMSAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQLIQOF	593
Db	481	ALVSQSALFSKGLAEPKMSAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQLIQOF	540
QY	594	AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLMDSG	653
Db	541	AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLMDSG	600
QY	654	RRFLLGTIPILIASLVILVSNLIDLGLAHALLSTVSVIVYFCCFVMGFGPIPNILCA	713
Db	601	RRFLLGTIPILIASLVILVSNLIDLGLAHALLSTVSVIVYFCCFVMGFGPIPNILCA	660
QY	714	EIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNATIGLAVFSIYAVVCLISFVFLK	773
Db	661	EIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNATIGLAVFSIYAVVCLISFVFLK	720
QY	774	VPETKMPLEVITEFFAVGAKQAATA 800	
Db	721	VPETKMPLEVITEFFAVGAKQAATA 747	

RESULT 3
AAU97204 standard; Protein; 737 AA.
XX AAU97204;
XX 27-AUG-2002 (first entry)
DE Soybean Arabidopsis thaliana-like sugar transport protein.
XX Soybean; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant.

XX Glycine max.
OS US6383776-B1.
XX 07-MAY-2002.
XX 14-APR-1999; 99US-0291922.
XX 24-APR-1998; 98US-083044P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2002-451386/48.
XX N-PSDB; ABK51965.
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX Claim 1; Fig 1; 54pp; English.

CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC represents a soybean Arabidopsis thaliana-like sugar transport protein.

XX Sequence 737 AA;
SQ Query Match 65.6%; Score 2674; DB 23; Length 737;
Best Local Similarity 69.9%; Pred. No. 2.7e-249;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY	56	MSGAVLVAIVASIGNLQGMWDNATIAAVALYIKKEFQONEPTVEGLIVSMGLIGATIVT	115
Db	1	MGAVMVAIAASIGNLQGMWDNATIGANGYIKKDALGT--TMERLVWGMGLIGATIVT	58
QY	116	TFSGPLSDSIGRRPMLILSSILYFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY	175
Db	59	TCSGPIADWLGRPRMILISSVLYFLGLVLMWSPNVYVLLARLIDGFGIGLAVTLVPLY	118
QY	176	ISEIAPSEIRGLNLTLFQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLFFFG	235
Db	119	ISETAPSEIRGLNLTLFQFSGSGGMFLSYCMVFGMSLSPAPSWRIMLGVLAIPLFFFG	178
QY	236	TIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGELSLLEGLVGGDTSEIYYIIGP	295
Db	179	TIFYLPESPRWLVSKGMAEAKVLOKLRGREDVSGEMALLVEGLIGGDTSEIYYIIGP	238
QY	296	ATEAADLVTDGKEQITLYGPEEGQSWIARPSKPIMLGSLVSLASRHGSMVNOQSVPLM	355
Db	239	ADDVADGHEHATEKXIRLYGSQAGLSWLSKPYTGQSSIG---LASHHGSLINQSMPLM	294
QY	356	DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEYA	413
Db	295	DPLVTLFGSIHEKLPETGARSGMRSTLFPNFGSMFSTAEPHAKIEQWDEESLOREREDYM	354
QY	414	SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRGSAISMROTLIGEG-GDGVSTSD	471
Db	355	SDATRGSDSDNLHSPILSRQATGAEGKDIVHHGRGSAISMROTLIGEG-GDGVSTSD	413
QY	472	ICGGWOLAMKWEKEGNGRKEGFKRVYLHOGVPGSRGSIIVSLPGGGDVFESEFVH	531
Db	414	ICGGWOLAMKWTDK-GEDEKQOGGFKRIRYLHEBGVSASRRGSIIVSLPGEG-----EFVQ	466

OS Arabidopsis thaliana.
XX
XX WO200210210-A2.
XX
XX 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX (FARB) BAYER AG.
XX
XX Tietjen K, Weidler M;
XX
XX WPI; 2002-269010/31.
XX
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX
XX Claim 5; SEQ ID NO 2463; 261pp + Sequence Listing; English.
XX
XX
XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
XX
SQ Sequence 729 AA;

Query Match 63.4%; Score 2584; DB 23; Length 729;
Best Local Similarity 68.8%; Pred. No. 1.4e-240;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
QY 56 MSGAVLVAIVASIGNLQGWMDNATIAAVALYIKKEFQIQNEPTVEGLIVSMSLIGATIVT 115
DB 1 MSGAVLVAIAAAGNLQGWMDNATIAAGVALYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGGLSDSIGRRPMLILSSILVFSSGLIMLWSPNVVYLLARFVDGFGIGLAVTLVPLY 175
DB 61 TCSGGVADWLGRRPMLILSSILVFSSGLVLMWSPNVVYLLGRLLDGFVGLVTLVPLY 120
QY 176 ISEIAPSEIRGLINTLPQFSGGGMLSYCMVFGMSLSPDPWRIMLGVALIPSLFFFG 235
DB 121 ISETAPPEIRGLINTLPQFTSGGMLSYCMVFGMSLSPDPWRIMLGVALIPSLVFFFL 180
QY 236 TIFYLPESPRWLVSXGMAEAKVLOKLRGKDVSGELSLLEGLVEVGDTISIEYITGP 295
DB 181 TVFFLPESPRWLVSXGMAEAKVLOKLRGKDVSGELSLLEGLVEVGDTISIEYITGP 240
QY 296 ATEAADLVLTGDKQITLYGPERQSWIARPSKGPIMLGVALISLRHGS-MVNQSVPL 354
DB 241 ADEVTDHDAVDKQIKLYGAEGLSWVARPVKG---GSTMSVLSRHGSTMSRQGS 296
QY 355 MDPIVTLFGSVHNMMPQAGGSMSTLPNFGSMFSVTDQAKNEQWDEENLHRDDEEYAS 414
DB 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPHFGSMFSVGNQPRHEDWDEENLVGEGDYP 355
QY 415 DGAGGDYEDNLHSPILSRQATGAEKDIVHGHGSAALSMRQTLLEGEGDVSSTDIG 474
DB 356 D-HGDDSEDDLHSPILSRQATGAEKDIVHGHGSAALSMRQTLLEGEGDVSSTDIG 413
QY 475 GWOGLAMKSEKEGNGRKEGFKRYVYLHQEGVPGSRGSIYLPGGGVDFEGSEFVHAA 534
DB 414 GWOVAMKTEREDESQKE-----EGFPGSRGSIYLPGGDGTGE-ADFVQASA 462
QY 535 LVSGSALFSKGLAPRMSDAAMVHSEVAAGSRWKDLFEFGVRALLVGVGIQILQQA 594

DB 463 LVSGPALYSKDLKEHTIGPAMVHPSE-TTKGSIWHDLDHPGVKRALVGVGIQILQQS 521
QY 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGFAMLMDLSGR 654
DB 522 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGFAMLMDLSGR 581
QY 655 RFLLLGTIPILIASLVIVSNLIDGLTAHALISTVSVIVYFCCFVMGFGPIPNILCAE 714
DB 582 RFLLLGTIPILIASLVIVSNLIDGLTAHALISTVSVIVYFCCFVMGFGPIPNILCAE 641
QY 715 IFPTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAGLAVGESIYAVVCLISFVFFLVK 774
DB 642 IFPTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAGLAVGESIYAVVCLISFVFFLVK 701
QY 775 PETKMPLEVITEFFAVGAKQAAA 798
DB 702 PETKMPLEVITEFFAVGAKQAAA 725

RESULT 6
ABB91078
ID ABB91078 standard; Protein; 734 AA.
XX
XX ABB91078;
XX
XX 31-MAY-2002 (first entry)
XX
XX Herbicidally active polypeptide SEQ ID NO 289.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX
XX Arabidopsis thaliana.
XX
XX WO200210210-A2.
XX
XX 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX (FARB) BAYER AG.
XX
XX Tietjen K, Weidler M;
XX
XX WPI; 2002-269010/31.
XX
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX
XX Claim 5; SEQ ID NO 289; 261pp + Sequence Listing; English.
XX
XX
XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
XX
SQ Sequence 734 AA;

Query Match 55.8%; Score 2275; DB 23; Length 734;
Best Local Similarity 61.2%; Pred. No. 1.2e-210;
Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;
QY 56 MSGAVLVAIVASIGNLQGWMDNATIAAVALYIKKEFQIQNEPT-VEGLIVSMSLIGATIV 114
DB 1 MSGAVLVAIAATIGNFLQGWMDNATIAAGAMVYINKDLNL---PTSVQGLVAVMSLIGATVI 57

QY	115	TFPSGPLSDSIGRRPMLILSSILYFSGGLIMLSPNVYLLLARFVDGFGIGLAVTLVPL	174
Db	58	TTCSGPISDMLGRRPMLILSSVMYFVCGGLIMLSPNVYLLCFARLLNGFAGLAVTLVPL	117
QY	175	YISEIAPSEIRGLINTLPQFSGSGGMFLSYCNVFGMSLSPDPWRIMLGVALPSLFFFG	234
Db	118	YISETAPEIRGQLNTLPQFLGSGGMFLSYCNVFTMSLSDSPSWRAMLGVALSIPSLYLF	177
QY	235	LTIFYLPESPRLVSKGRMAEAKKVLQKLRGDVSGELSLLEGLFVGGDTSTIEYIIG	294
Db	178	LTVFYLPESPRLVSKGRMDEAKRVLQQLCGREDVTIDEMALLVEGLDIGEKTIMEDLLVT	237
QY	295	PATEAADLVTDGKE-QITLYGPEEGQSWTARPSKGPIMLGSVLSLASRHGSVMVQSV	353
Db	238	LEHGEDDTLETVEDEGQRLYGTHENQSYLARPFEEQ---NSSLGLRSRHGSLANQSMI	294
QY	354	LMDDPIVTLFGSVHENMPQAGGSMRSTLPNFGSMFVTDQ--HAKNEQWD--BENLHRD	408
Db	295	LKDPVLNLFGLHEKMPGAGNTRSGIFPHFGSMFTTADAPHGKPAHWEKDIESHYNKD	354
QY	409	DEEYASDAGGDYED--NLHSPLSRQATGAECKDIVHGHRSALSMRQTLLEGGD	465
Db	355	NDDYATDDGAGDDDDSDNDLRSPLMRSQTSMQ-KIMI PHTSGSTLSMRHSTLMQ-GN	412
QY	466	GVSSTDIGGQWLAMKMWSEKEGENGRKEGFKRYLIHQEGVPGSRGSI VSLPGGDVFE	525
Db	413	GESSMGIGGGWHMGYRYENDE-----YKRYLIKEDGAE-SRRGSIISIFGGPD--G	460
QY	526	GSEFVHAALVSQSALFSKGLAEPMSDAMVHPSEVAAKSGRWKDLFEPGYRALLVGV	585
Db	461	GGSYIHASALVSRSVLGPKS----VHGSAVDPKIIAASGPLMSALLEPGYKRALVGV	515
QY	586	GIQILQFAGINGVLYTTPQILBEQAVAVILSKFGLSSASASILSSLTLLMLPCIGFA	645
Db	516	GIQILQFSGINGVLYTTPQILERAGVDILSSLGSSISASFLISGLTLLMLPAI VVA	575
QY	646	MLMDLSGRFLLGTIPILIASLVILVSNLIDLGLAHALLSTVSVIYVFCFCVMGFG	705
Db	576	MRMDVSGRSLLLWTIPVLIVSLVAVISELHISKVNNAALSTGCVVLYFCFCVMGYG	635
QY	706	PIPNILCAEIEPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAIGLAVFSIYAVCLI	765
Db	636	PIPNILCSEIEPTRVRGLCIAICAMFWIGDIIVTYSLPVLLSSIGLVGVFSIYAAVCVI	695
QY	766	SFVFVFLKVPETKGMPLVITEFFAVGAKQAATA 800	
Db	696	SWIFVYMKVPETKGMPLVITDYFAFGA-QAQASA 729	

RESULT 7	
ID	ABB92622 standard; Protein; 729 AA.
XX	
AC	ABB92622;
XX	
DT	31-MAY-2002 (first entry)
XX	
DE	Herbicidally active polypeptide SEQ ID NO 1833.
XX	
KW	Herbicidal; plant; agriculture; herbicide.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200210210-A2.
XX	
PD	07-FEB-2002.
XX	
XX	
PF	28-AUG-2001; 2001WO-EP09892.
XX	
XX	
PR	28-AUG-2001; 2001WO-EP09892.
XX	
PA	(FARB) BAYER AG.

XX Tietjen K, Weidler M;
PI
XX WPI; 2002-269010/31.
DR
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 1833; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
XX useful as herbicides.
XX
SQ Sequence 729 AA;

[illegible]

Db 627 FGAI PNILCS EIFTSVRG LITICALT FWICDI IYTYTL PVM LKSI GAGVFIYA IVC 686
QY 764 LISFV FVLKVPETKGM PLEVITEFFAVGAKQA AKA 800
Db 687 AVAMVFVY LKVPETKGM PLEVISEFFSVGAKQDAAA 723

RESULT 8
AAU97205

ID AAU97205 standard; Protein; 486 AA.

AC AAU97205;

DT 27-AUG-2002 (first entry)

DE Portion of a soybean sugar transport protein encoded by ssl.pk0022.fl.

KW Soybean; Arabidopsis thaliana-like sugar transport protein;
KM carbohydrate transport; grain filling; annual field crop;
XX plant.

OS Glycine max.

PN US6383776-B1.

PD 07-MAY-2002.

PF 14-APR-1999; 99US-0291922.

PR 24-APR-1998; 98US-083044P.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

DR WPI; 2002-451386/48.

DR N-PSDB; ABK51966.

PT New nucleic acid encoding plant sugar-transport proteins, useful for

PS Claim 1; Fig 1; 54pp; English.

CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC represents a portion of a soybean Arabidopsis thaliana-like sugar
CC transport protein.

SO Sequence 486 AA;

Query Match 41.0%; Score 1670.5; DB 23; Length 486;
Best Local Similarity 68.1%; Pred. No. 2.2e-152;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

QY 308 DKEQITLYGPEEGSWIARPSKPIMLGSLSLARHSGSMVNQSVPLMDPIVTLFGSVHE 367

Db 5 EKDQIKLYGPEQGQSWVARPVAGPNSVG---LVSRKGSMA NPS-SLVDPVLTLYFGSVHE 59

QY 368 NMPQAGSGMRSTLFPNFGSMFSVTDQAKNEQWDEENLHRDDEBYASDAGADYEDNLHS 427

Db 60 KLPE TG----STLFPHFGSMFSVGNQPRNEDWDESLAREGDDYVD--AGSDDDLQOS 113

QY 428 PLLSRQATGAEKSDIVHHGHSALSMRQTL--GEGGDVGSSTDIGGQQLAWKMSKE 486

Db 114 PLISRQ TSLD-KDIPPHA NSNLA-SMRQSLH GNSGEPTGSTGIGGQQLAWKMSERE 171

QY 487 GENGRKEG GFKRYLIHQEGVP GSRGSI VSLPGGDVFE GSEFVHAALYSQALFSKGL 546

Db 172 GPDGKEG GFKRIYLHQDGSGSRRGSVSLP-GGDLPTDSEVQAALVSPALYNEDL 230

QY 547 AEP RMDAAMVHPSEVA AKSRWKDLFEPGVRALLVGVGIQI LQOFAGINGLYTPQI 606

Db 231 MRQRPVGAMIHPS ETIAKGPSWSDLFEPGVKHALIVGVMOI LQOFSGINGVLYTPQI 290

QY 607 LEQAGVAVILSKRG LSSASASILISSLTLLMLPCIGFAMLLMDL SGRPLLGTIPILI 666

Db 291 LEQAGVGYLLSLGL GSTSSFLISA VTLMLPCIAIAMRLMDISGR TLLISTIPVLI 350

QY 667 ASLVILVSNLIDGLTAHALSTVSIVYFC CFVMGFGPIPNILCAEIPTRVRLCIA 726

Db 351 AALLILVLSLVDLGSTANASISTISVIYFC FVMGFGPIPNILCAEIPTRVRLCIA 410

QY 727 ICAFTFWIGDIIVTYSLPVMLNAGLGVSIYAVVCLISFVFLKVPETKGM PLEVIT 786

Db 411 ICALTFWICDIIVTYTL PVM LNSVGLAGVGIYAVVCFIAWVFVFLKVPETKGM PLEVIT 470

QY 787 EFFAVGAKQ 795

Db 471 EFFSVGAKQ 479

RESULT 9

ID ABU08330 standard; Protein; 486 AA.

AC ABU08330;

DT 29-MAY-2003 (first entry)

DE Soybean sugar transport protein #2.

KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;

KM plant sugar transport protein; carbohydrate transport; soybean;

OS Glycine max.

PN US2002178468-A1.

PD 28-NOV-2002.

PF 17-JAN-2002; 2002US-0051902.

PR 24-APR-1998; 98US-083044P.

PR 14-APR-1999; 99US-0291922.

PA (ALLE/) ALLEN S M.

PA (HITZ/) HITZ W D.

PA (KINN/) KINNEY A J.

PA (TING/) TINGEY S V.

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

DR WPI; 2003-340957/32.

DR N-PSDB; ABX93202.

PT Novel plant sugar transport proteins and nucleic acid encoding the

PT protein useful for producing transgenic plants having altered levels of

PT sugar transport protein -

PS Claim 5; Fig 1; 56pp; English.

CC The present invention relates to the isolation of Arabidopsis

CC thaliana-like or Beta vulgaris-like sugar transport proteins, and

CC the polynucleotide sequences encoding them. The plant sugar transport

CC proteins of the invention have been isolated from corn, rice, soybean,

CC and wheat. The polypeptides of the invention may be used for altering

the level of expression of a sugar transport protein in a host cell, by transforming a host cell with a chimeric construct encoding all, or a portion of the sugar transport protein, in sense or antisense orientation. Particularly, the polypeptides may provide a means to control carbohydrate transport and distribution in plants. ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar transport proteins.

SQ Sequence 486 AA;

Query Match	41.0%;	Score 1670.5;	DB 24;	Length 486;
Best Local Similarity	68.1%;	Pred. No. 2.2e-152;		
Matches 333;	Conservative 57;	Mismatches 84;	Indels 15;	Gaps 8;

QY	308	DKEJTLVGPBEEGOSWIARPSKPIMLGSLVSLASRHGSMVNOSVPLMDPIVTLFGSVHE	367
Db	5	EKDQIKLYGPEQGOSWVARPVAGPNSVG---LVSRKGSMANPS-SLVDPVLVTLFGSVHE	59
QY	368	NMPDAGGSMRSTLFPNFGSMFSVTDQHAKNQEWDEENLHRDDEEYASDGAGDYEDNLHS	427
Db	60	KLPEFG-----STLFPHFGSMFSVGNQPRNEDWDEESLAREGDDYVD--AGSDNLIQS	113
QY	428	PLLSROATGAEGKDIVHGHGRSALSMRQTLI-GEKGDVSTDIGGQOLAWKMSKE	486
Db	114	PLISRQTTSLD-KDIPPHASNLA-SMRQSLHGNSEPTGSTGIGGQOLAWKMSERE	171
QY	487	GENGRKEGCFKRVYLHOEGVPGSRGSIIVSLPGGDDVFESEFVHAALVSOALFSKGL	546
Db	172	GPDSKEGCFKRIYLIHQDGGSGSRGSIIVSLP-GGDLPTDSEVVQAALVSOALVYNEDL	230
QY	547	AEPNMSDAMVHPSEVAAGSRWKDLFEPGYRALLVGVGIILOQFAGINGVLYTTPQI	606
Db	231	MROGPVGPAMIHPSSETIAKGPSWDLFEPGVKHALIVGVMOILQOFSGINGVLYTTPQI	290
QY	607	LEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSGRFLLCTPIILI	666
Db	291	LEQAGVAVILSSGLGSTSSSFLISAVTLLMLPCIAIAMRLMDISGRRTLLSTIPVLI	350
QY	667	ASLVLVSNLIDLGLTALHALLSTSVIVYFCCFVMGFGPIPNILCAEIPFTRVGLCIA	726
Db	351	AALLVLVGLSLVDLGSTANASISTISIVYFCCFVMGFGPIPNILCAEIPFTRVGLCIA	410
QY	727	ICAFPFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLKVPETKMPLEVIT	786
Db	411	ICALTFWICDIIIVTTLPMNLNSVGLAGVFGIYAVVCFIAMVFFLKVPEIKMPLEVII	470
QY	787	EEFFAVGAKQ 795	
Db	471	EEFFSVGAKQ 479	

XX	14-APR-1999;	99US-0291922.
XX	PF	
XX	PD	
XX	07-MAY-2002.	
XX	US6383776-B1.	
XX	PN	
XX	OS	
XX	Triticum aestivum.	
KW	carbohydrate	
KW	wheat; Arabidopsis thaliana-like sugar transport protein;	
XX	plant.	
XX	Portion of a	
XX	27-AUG-2002	
XX	DT	
XX	AC	
XX	AAU97207;	
XX	ID	
AAU97207	AAU97207 standard; Protein; 345 AA.	
RESULT 10		

PR 24-APR-1998; 98US-083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-451386/48.
DR N-PSDB; ABK51968.
XX
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
PT
XX
PS Claim 1; Fig 1; 54pp; English.

DR WPI; 2002-451386/48.
DR N-PSDB; ABK51968.

PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
XX Claim 1, Fig 1, 54pp; English.
PS
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC represents a portion of a wheat Arabidopsis thaliana-like sugar
CC transport protein.

SQ **Sequence** **345** **AA;**

Query Match	34.9%	Score 1424;	DB 23;	Length 345;
Best Local Similarity	82.9%	Pred. No. 9.2e-129;		
Matches 281; Conservative	24;	Mismatches 32;	Indels 2;	Gaps 22;

QY	462	EGGDVSVSTDIGGQWOLAMKWSSEKENGNGRKEGGFKRVYLHQEGVPGSRGSI VSLPGGG	521
Db	4	EGGEAVSVSTGIGGGQWLAMKWSERQGEDGKKEGGFKR IY LHQEGVADSRGRSVSLPGGG	63
QY	522	DVFE -GSEFVHAALVSOALVPSKGLAEPRM -SDAAMVHPSEVAAGKSRWDLFEPGVRR	579
Db	64	DATQGGSGFIHAALVSHSALYSKDLMERMAAGPAMIHPLAAPKGSIMKDLFEPGVRR	123
QY	580	ALLVGVGIOILQOFAGINGVLYTTPQILBQAGVAVILSKFGLSSASASILLISLTTLLML	639
Db	124	ALFVGVGIOQLQOFAGINGVLYTTPQILBQAGVAVLLSNLGLSSASASILLISLTTLLML	183
QY	640	PCIGFAMLLMDLSGRFLILGTIPILLASLVILVSNLIDLGTLAHALLSTVSVIVYFCC	699
Db	184	PSIGVAMRLMDISGRFLILGTIPILLASLVILGVNVINLSTVPHAVLSTVSVIVYFCC	243
QY	700	FVMGFGPIPNILCAEIPFTRVRCICIAICAFTFWIGDIIVTYSLPVMLNAIGLAGVFSIY	759
Db	244	FVMGFGPIPNILCAEIPFTRVRCVCAICALTFWICDIIVTYSLPVMLNAIGLAGVFGIY	303
QY	760	AVVCLISFVFLKVPETKGMPLLEVITBFFAVGAKOAAA	798
Db	304	AVVCCIAFVEVYLKVPETKGMPLLEVITBFFAVGAKOAAA	342

RESULT 11	
ABU08332	
ID	ABU08332 standard; Protein; 345 AA.
XX	
AC	ABU08332;
XX	
DT	29-MAY-2003 (first entry)
XX	
DE	
XX	Wheat sugar transport protein #2.
KW	Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW	plant sugar transport protein; carbohydrate transport; soybean;
KW	carbohydrate distribution; plant.
XX	
OS	Triticum aestivum.

XX US2002178468-A1.
PN
XX
XX 28-NOV-2002.
PD
XX
XX 17-JAN-2002; 2002US-0051902.
PF
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALIE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR N-PSDB; ABX93204.
XX
XX Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
XX
PS Claim 5; Fig 1; 56pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
CC transport proteins.
XX
SQ Sequence 345 AA;

Query Match 34.9%; Score 1424; DB 24; length 345;
Best Local Similarity 82.9%; Pred. No. 9.2e-129;
Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;

QY 462 EGGDGVSTDIGGQWLAWKWEKEGNGRKEGFKRVYLHOGVPGSRGSIIVSLPGCG 521
Db 4 EGGEAVSTSTIGGGQWLAWKWEKEGDEGKKEGFKRIYLHOGVADSRGVSIVSLPGGG 63

QY 522 DVFE-GSEFVHAALVSSALFSKGLAERPM-SDAAMVHPSEVAAKGSRWKDLFEFGVRR 579
Db 64 DATCGSGFIHAALVSHSALYKDLMEERMAAGPAMIHPLEAAPKGSIMKDLFEFGVRR 123

QY 580 ALVGVGIIQILQQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASISLITLML 639
Db 124 ALFVGVGIQMLQQFAGINGVLYTTPQILEQAGVAVILSNLGLSSASASISLITLML 183

QY 640 PCIGFAMLLMDLSGRRFLLGTIPILIASLVILVNSNLDIGTLAALLSTVSVIYFCC 699
Db 184 PSIGVAMRLMDISGRFLLGTIPILIASLVILVNSNLDIGTLAALLSTVSVIYFCC 243

QY 700 FVMGFGPIPNILCAEIPFTRVRGLCIACAFTFWIGDIIVTYSLPVMLNAIGLAGVFSY 759
Db 244 FVMGFGPIPNILCAEIPFTRVRGLCIACAFTFWIGDIIVTYSLPVMLNAIGLAGVFSY 303

QY 760 AVVCLISFVFLKVPETKGMPLVITEFFAVGAKQAAA 798
Db 304 AVVCCIAFVFLKVPETKGMPLVITEFFAVGAKQAAA 342

RESULT 12
AAU97208 standard; Protein; 228 AA.
ID

XX AAU97208;
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX Portion of a wheat sugar transport protein encoded by wreln.pk0006.b4.
DE
XX
XX Wheat; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant.
XX
XX Triticum aestivum.
OS
XX
XX US6383776-B1.
PN
XX
XX 07-MAY-2002.
PD
XX
XX 14-APR-1999; 99US-0291922.
PF
XX
XX 24-APR-1998; 98US-083044P.
PR
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-451386/48.
DR N-PSDB; ABK51969.
XX
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 1; Fig 1; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC represents a portion of a wheat Arabidopsis thaliana-like sugar
CC transport protein.
XX
SQ Sequence 228 AA;

Query Match 23.1%; Score 943; DB 23; length 228;
Best Local Similarity 80.6%; Pred. No. 1.8e-82;
Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 574 EPGVRALLVGVGIQILQQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASISL 633
Db 1 EPGVKHALFVGIGLQILQQFAGINGVLYTTPQILEQAGVAVILSNLGLSSASISL 60

QY 634 TTLMLPCIGFAMLLMDLSGRRFLLGTIPILIASLVILVNSNLDIGTLAALLSTVSV 693
Db 61 TTLMLPSIGIAMRLMDMSGRRFLLSTIPVILVALVILVNLVDGTMVHALSTISV 120

QY 694 IVYFCCFVMGFGPIPNILCAEIPFTRVRGLCIACAFTFWIGDIIVTYSLPVMLNAIGLA 753
Db 121 IVYFCFVMGFGPIPNILCAEIPFTRVRGLCIACAFTFWIGDIIVTYSLPVMLNAIGLA 180

QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLVITEFFAVGAKQ 795
Db 181 GVFGIYAVVCLAFVFLKVPETKGMPLVITEFFAVGAKQ 222

RESULT 13
ABU08333 standard; Protein; 228 AA.
ID

AC ABU08333;
XX
DT 29-MAY-2003 (first entry)
XX
DE Wheat sugar transport protein #3.
XX
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX
OS Triticum aestivum.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002
XX
PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2003-340957/32.
DR N-PSDB; ABX93205.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
XX
XX Claim 5; Fig 1; 56pp; English.
PS
XX The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
CC transport proteins.
XX
SQ Sequence 228 AA;
Query Match 23.1%; Score 943; DB 24; Length 228;
Best Local Similarity 80.6%; Pred. No. 1.8e-82;
Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;
QY 574 EPGYRALLVGVGIQILQGFAGINGVLYTPQILEQAGVAVILSKFGLSSASASILISL 633
Db 1 EPGVXHALFVGIGLQILQGFAGINGVLYTPQILEQAGVAVILSNIGLSSSSASILISAL 60
QY 634 TTLMLPCIGFAMLMDLGRRPLLLGTIPILIASIVLVNSNLDLGLAHALSTVSV 693
Db 61 TTLMLPSIGIAMRLMDSGRRPLLLSTIPVLIVALAVLVNVDVGTMVHAALSTISV 120
QY 694 IVYFCCFVMGFGPIPNILCAEIPFTRVRLCIAICAFTFWIGDIIVTYSLPVMLNAIGLA 753
Db 121 IVYFCFVMGFGPIPNILCAEIPFTRVRLCIAICAFTFWIGDIIVTYSLPVMLNAIGLA 180
QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLLEVITEFFAVGAKQ 795
Db 181 GVFGIYAVVCLISFVFLKVPETKGMPLLEVITEFFAVGAKQ 222

RESULT 14
AAU97203
ID AAU97203 standard; Protein; 131 AA.
XX
XX AAU97203;
AC
XX
DT 27-AUG-2002 (first entry)
XX
DE Portion of a rice sugar transport protein encoded by r186.pk0003.d5.
XX
KW Rice; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant.
XX
OS Oryza sativa.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-0291922.
XX
PR 24-APR-1998; 98US-083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-451386/48.
DR N-PSDB; ABK51964.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
XX Claim 1; Fig 1; 54pp; English.
PS
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC represents a portion of a rice Arabidopsis thaliana-like sugar
CC transport protein.
XX
SQ Sequence 131 AA;
Query Match 13.8%; Score 562; DB 23; Length 131;
Best Local Similarity 80.5%; Pred. No. 5.8e-46;
Matches 103; Conservative 17; Mismatches 8; Indels 0; Gaps 0;
QY 669 LVILVNSNLDLGLAALLSTVSVIVYFCCFVMGFGPIPNILCAEIPFTRVRLCIAIC 728
Db 1 VLTLLVNLIDVGTWVHASLSTVSVILYFCFVMGFGPIPNILCAEIPFTRVRLCIAIC 60
QY 729 AFTFWIGDIIVTYSLPVMLNAIGLAVFSIYAVVCLISFVFLKVPETKGMPLLEVITEF 788
Db 61 AFTFWIGDIIVTYSLPVMLNAIGLAVFSIYAVVCLISFVFLKVPETKGMPLLEVITEF 120
QY 789 FAVGAKQA 796
Db 121 FSVGAKQA 128
RESULT 15
ABU08328
ID ABU08328 standard; Protein; 131 AA.
XX
XX ABU08328;

XX		29-MAY-2003	(first entry)
DT		Rice sugar transport protein #2.	
XX			
DE		Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;	
KW		plant sugar transport protein; carbohydrate transport; soybean;	
KM		carbohydrate distribution; plant.	
XX			
OS		Oryza sativa.	
XX			
PN		US2002178468-A1.	
XX			
PD		28-NOV-2002.	
XX			
PJ		17-JAN-2002; 2002US-0051902.	
XX			
PF		24-APR-1998; 98US-083044P.	
PR		14-APR-1999; 99US-0291922.	
XX			
PA		(ALLE/) ALLEN S M.	
PA		(HITZ/) HITZ W D.	
PA		(KINN/) KINNEY A J.	
PA		(TING/) TINGEY S V.	
PI			
PI		Allen SM, Hitz WD, Kinney AJ, Tingey SV;	
DR		WPI; 2003-340957/32.	
XX		N-PSDB; ABX93200.	
PT			
PT		Novel plant sugar transport proteins and nucleic acid encoding the	
PT		sugar transport protein -	
XX			
PS		Claim 5; Fig 1; 56pp; English.	
CC			
CC		The present invention relates to the isolation of Arabidopsis	
CC		thaliana-like or Beta vulgaris-like sugar transport proteins, and	
CC		the polynucleotide sequences encoding them. The plant sugar transport	
CC		proteins of the invention have been isolated from corn, rice, soybean,	
CC		and wheat. The polypeptides of the invention may be used for altering	
CC		the level of expression of a sugar transport protein in a host cell,	
CC		by transforming a host cell with a chimeric construct encoding all,	
CC		or a portion of the sugar transport protein, in sense or antisense	
CC		orientation. Particularly, the polypeptides may provide a means to	
CC		control carbohydrate transport and distribution in plants.	
CC		ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar	
CC		transport proteins.	
SQ			
SQ		Sequence 131 AA;	
Query Match		13.8%; Score 562; DB 24; Length 131;	
Best Local Similarity		80.5%; Pred. No. 5.8e-46;	
Matches 103; Conservative		17; Mismatches 8; Indels 0; Gaps 0;	
QY		669 LVILVNSNLI DLGLTAHALISTVSIVIVFCGFMVGFGPIPNI LCAEIFPTRVRGLCIAIC	728
Db		1 VLTLILVNILDVGTMTVMHASLSTVSILLYFCFVMGFGPIPNI LCAEIFPTRVRGICIAIC	60
QY		729 AFTFWIGDIIIVTYSLPVMLNAIGLAGVESIYAVVCLISFVFVFLKVPETKGMPLVEITTF	788
Db		61 ALTFWIGDIIIVTYTLPVMLNAIGLAGVFGIYAVVCILAFLFVFMKVPETKGMPLVEITTF	120
QY		789 FAVGAKQA 796	
Db		121 FSVGAKQA 128	

Search completed: January 5, 2004, 18:55:14
Job time : 78 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:54:10 ; Search time 30 Seconds
(without alignments)
1128.290 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQPTFTPLDRR.....PLEVITTEFFAVGAKQAATA 800

Scoring table: BL06UM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	3517	86.3	747	US-09-291-922-2	Sequence 2, Appli
2	2674	65.6	737	US-09-291-922-8	Sequence 8, Appli
3	2584	63.4	729	US-09-291-922-29	Sequence 29, Appli
4	1670.5	41.0	486	US-09-291-922-10	Sequence 10, Appli
5	1424	34.9	345	US-09-291-922-14	Sequence 14, Appli
6	943	23.1	228	US-09-291-922-16	Sequence 16, Appli
7	562	13.8	131	US-09-291-922-6	Sequence 6, Appli
8	496	12.2	523	US-09-291-922-24	Sequence 24, Appli
9	489.5	12.0	513	US-09-291-922-20	Sequence 20, Appli
10	489	12.0	529	US-09-291-922-28	Sequence 28, Appli
11	488	12.0	549	US-09-291-922-30	Sequence 30, Appli
12	485	11.9	510	US-09-291-922-22	Sequence 22, Appli
13	463.5	11.4	539	US-09-291-922-26	Sequence 26, Appli
14	421.5	10.3	584	US-08-928-692-13	Sequence 13, Appli
15	421.5	10.3	584	US-09-339-972-13	Sequence 13, Appli
16	384.5	9.4	534	US-09-031-392-4	Sequence 4, Appli
17	384.5	9.4	534	US-09-299-549-4	Sequence 4, Appli
18	384.5	9.4	534	US-09-610-417-4	Sequence 4, Appli
19	365	9.0	493	US-09-031-392-10	Sequence 10, Appli
20	365	9.0	493	US-09-299-549-10	Sequence 10, Appli
21	365	9.0	493	US-09-610-417-10	Sequence 10, Appli
22	358	8.8	488	US-08-928-692-10	Sequence 10, Appli
23	358	8.8	488	US-09-339-972-10	Sequence 10, Appli
24	346	8.5	524	US-08-928-692-12	Sequence 12, Appli
25	346	8.5	524	US-09-339-972-12	Sequence 12, Appli
26	344	8.4	509	US-09-031-392-6	Sequence 6, Appli
27	344	8.4	509	US-09-299-549-6	Sequence 6, Appli

28	344	8.4	509	4	US-09-610-417-6	Sequence 6, Appli
29	331	8.1	492	2	US-08-355-844-3	Sequence 3, Appli
30	331	8.1	492	5	PCT-US95-16126-3	Sequence 3, Appli
31	330	8.1	488	2	US-08-928-692-11	Sequence 11, Appli
32	330	8.1	488	4	US-09-339-972-11	Sequence 11, Appli
33	310	7.6	494	2	US-09-031-392-5	Sequence 5, Appli
34	310	7.6	494	3	US-09-299-549-5	Sequence 5, Appli
35	310	7.6	494	4	US-09-610-417-5	Sequence 5, Appli
36	306.5	7.5	500	2	US-09-299-549-7	Sequence 7, Appli
37	306.5	7.5	500	3	US-09-299-549-7	Sequence 7, Appli
38	306.5	7.5	500	4	US-09-610-417-7	Sequence 7, Appli
39	293	7.2	383	2	US-09-031-392-3	Sequence 3, Appli
40	293	7.2	383	3	US-09-299-549-3	Sequence 3, Appli
41	293	7.2	383	4	US-09-610-417-3	Sequence 3, Appli
42	275.5	6.8	323	4	US-09-134-001C-4635	Sequence 4635, Ap
43	265.5	6.5	475	4	US-09-328-352-5706	Sequence 5706, Ap
44	256.5	6.3	131	4	US-09-291-922-4	Sequence 4, Appli
45	256	6.3	305	4	US-09-107-532A-4215	Sequence 4215, Ap

ALIGNMENTS

Applicant
Not a download

RESULT 1
US-09-291-922-2
Sequence 2, Application US/09291922
Patent No. 6382226
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 747
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (129)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (133)..(134)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (144)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (178)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (207)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (218)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (220)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (236)
US-09-291-922-2
Query Match 86.3%; Score 3517; DB 4; Length 747;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;


```
QY 56 MSGAVLVAIVASIGNLLOGWDNATIAAVALYIKKEFOLONEPTEGLIVSMISLIGATVIT 115
Db 1 MGSAVVAIAASIGNLLOGWDNATIAAGAVLYIKKEFNLOSEPLIEGLIVAMFLIGATVIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGGLIMLWSPNVYVLLARFVDFGFIGLAATLVPLY 175
Db 61 TSPGPRADCVGRPMLVASAVLYFVSGVLMWAPIVYILLARLIDGFIGLAATLVPLY 120
QY 176 ISEIAPSEIR-GLINTLPQFSG-SGGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFF 233
Db 121 ISETAPHRXSWGXXNTLPQFIGYCGMFLSYCMVFGMSLMPKPDWRIMLGVALIPSLXYF 180
QY 234 GLTIFYLPESPRWLVSCKRMAEAKVLOKLRKDVSGELSLLEGLVGGDTSEIETII 293
Db 181 GLTIFYLPESPRWLVSCKRMAEAKVYQRLRGREDVSXEXALLVEGLGVGKDTRIXEYII 240
QY 294 GPATEAADLVTDGKEQITLYGPBEGQSWIARPSKGPIMLGSVLSLASRHGSMVNOQVP 353
Db 241 GPATEAADLVTDGKEQITLYGPBEGQSWIARPSKGPIMLGSVLSLASRHGSMVNOQVP 300
QY 354 LMDPIVTLFGSVHENMPQAGSMSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYA 413
Db 301 LMDPIVTLFGSVHENMPQAGSMSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYA 360
QY 414 SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRSALSMBROTLLGEGGDVSSSTDIG 473
Db 361 SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRSALSMBROTLLGEGGDVSSSTDIG 420
QY 474 GGMQOLAMKWESEKEGNGRKEGFKRVYLIHQEGVPGSRGSIYSLPGGQDVFESEFVHA 533
Db 421 GGMQOLAMKWESEKEGNGRKEGFKRVYLIHQEGVPGSRGSIYSLPGGQDVFESEFVHA 480
QY 534 ALVSQALFSKGLAEPKMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVIQIQLQF 593
Db 481 ALVSQALFSKGLAEPKMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVIQIQLQF 540
QY 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSG 653
Db 541 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDSG 600
QY 654 RRFLLGTIPILIASLVLIVSNLIDGLTAHALLSTVSVIYFCCFVMGFGPIPNILCA 713
Db 601 RRFLLGTIPILIASLVLIVSNLIDGLTAHALLSTVSVIYFCCFVMGFGPIPNILCA 660
QY 714 EIFPTRVRGLCIAICFTFWIGDIIVTYSLPVMLNAIGLAVFSIYAVVCLISFVFEVLK 773
Db 661 EIFPTRVRGLCIAICFTFWIGDIIVTYSLPVMLNAIGLAVFSIYAVVCLISFVFEVLK 720
QY 774 VPETKGMPLLEVITEFFAVGAKQAAKA 800
Db 721 VPETKGMPLLEVITEFFAVGAKQAAKA 747
```

```
RESULT 2
US-09-291-922-8
; Sequence 8, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 737
```

```
; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-8
Query Match 65.6%; Score 2674; DB 4; Length 737;
Best Local Similarity 69.9%; Pred. No. 8.9e-256;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;
QY 56 MSGAVLVAIVASIGNLLOGWDNATIAAVALYIKKEFOLONEPTEGLIVSMISLIGATVIT 115
Db 1 MGSAVVAIAASIGNLLOGWDNATIAAGAVLYIKKEFNLOSEPLIEGLIVAMFLIGATVIT 58
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGGLIMLWSPNVYVLLARFVDFGFIGLAATLVPLY 175
Db 59 TCSGPIDWLGRRPMMIISVLYFLGGLVLMWSPNVYVLLARLIDGFIGLAATLVPLY 118
QY 176 ISEIAPSEIRGLINTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFF 235
Db 119 ISETAPSEIRGLINTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFF 178
QY 236 TIFYLPESPRWLVSCKRMAEAKVLOKLRKDVSGELSLLEGLVGGDTSEIETIIGP 295
Db 179 TIFYLPESPRWLVSCKRMAEAKVLOKLRGREDVSGEMALLVEGLIGGDTSEIETIIGP 238
QY 296 ATEAADLVTDGKEQITLYGPBEGQSWIARPSKGPIMLGSVLSLASRHGSMVNOQVP 355
Db 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPYTGQSSIG---LASHGSIINQSMPLM 294
QY 356 DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYA 413
Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLFPNFGSMFSTAEPHAKIEQWDEESLQREDYDM 354
QY 414 SDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRSALSMBROTLLGEG-GDGVSSSTD 471
Db 355 SDATRGSDSDNLHSPILSRQTTSLK-KDLPPRPSHGSIIGSMRHSLSMQSGEGGSGTG 413
QY 472 ICGGQOLAMKWESEKEGNGRKEGFKRVYLIHQEGVPGSRGSIYSLPGGQDVFESEFVH 531
Db 414 ICGGQOLAMKWTDK-GEDEKQOQGFKRITYLHEEGVSASRGSIVSIPGEG-----EFVQ 466
QY 532 AAALVSQALFSKGLAEPKMSDAMVHPSEVAAGSRWKDLFEPGVRALLVGVIQIQLQ 591
Db 467 AAALVSQALYSKELIDGHPVGPMVHPSETASKGPSWKALLEPGVKHALVGVIQIQLQ 526
QY 592 QFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDL 651
Db 527 QFSGINGVLYYTPQILEEAGVEVLLSDIGIGSESASFLLISFTTFLMLPCIGVAMKLMV 586
QY 652 SGRFLLGTIPILIASLVLIVSNLIDGLTAHALLSTVSVIYFCCFVMGFGPIPNIL 711
Db 587 SGRQQLLTTPILVLSLILVIGSLVNFVNAHAISTVCVVVYFCCFVMGFGPIPNIL 646
QY 712 CAEIFPTRVRGLCIAICFTFWIGDIIVTYSLPVMLNAIGLAVFSIYAVVCLISFVFEVL 771
Db 647 CSEIFPTRVRGLCIAICLVFWIGDIITYSLPVMLGSLGLGVFAIYAVVCFISWIFVF 706
QY 772 LKVPETKGMPLLEVITEFFAVGAKQAAA 798
Db 707 LKVPETKGMPLLEVITEFFAVGAKQAAAS 733
```

```
RESULT 3
US-09-291-922-29
; Sequence 29, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
```

```
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-291-922-29
```

```
Query Match      63.4%; Score 2584; DB 4; Length 729;
Best Local Similarity 68.8%; Pred. No. 7.3e-247;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
```

```
QY 56 MSGAVLVAIVASIGNLQGWDMATIAAALVYIKKEFOLQNEPTVEGLIVSMSLIGATIVT 115
Db 1 MSGAVLVAIAAVGNLLQGWDMATIAAVLYIKKEFNLESNPVEGLIVAMSLIGATLIT 60
QY 116 TFGSPLSDSIGRRPMLILSSILYFSGSLIMLMSPNVYVLLARFVDGFGIGLAVTVPLY 175
Db 61 TCSGGVADWLGRRPMLILSSILYFVGSVLMSPNVYVLLGRLLDGFVGLVTVLPY 120
QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPWRIMLGVLAIPLSFEGL 235
Db 121 ISETAPPEIRGLNLTLPQFTGSGMFLSYCMVFGMSLSPSPWRIMLGVLFIPLVFPFL 180
QY 236 TIFYLPESPRWLVSKGMAEAKVYLQKRGKDVSGELSLLEGLVGGDTSEIYYIGP 295
Db 181 TVFFLPESPRWLVSKGRLAKRVQLRGRDVSSEMALLEGLIGGETTIEEYIIGP 240
QY 296 ATEAADLVTDGDEKQITLYGPEEGSWIARPSKPIMLGSVLSLRHGS-MVNQSVPL 354
Db 241 ADEVTDHDHDAVDKDIKLYGAEEGLSWARPVKG---GSTMSVLSRHGSTMSRRQSL 296
QY 355 MDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSVTQDHAKNEQWDEENLHRDDEEYAS 414
Db 297 IDPLVTLFGSVHEKMPDT-GSMRSLFPHFGSMFSGVGNQPRHEDWDEENLVGEEDYPS 355
QY 415 DGAGGYEDNLHSPPLSRQATGAEGKDIVHGHRSALSMRQTLLEGGDVSTDTIG 474
Db 356 D-HGDDSEDDLHSPILSRQATSMF-KDMPHTAHGTLSTFRHGSQYQAGGEGAGSMGIG 413
QY 475 GWQLAWKSEKEGEGNGRKEGFKRYVLHQEGVPSRRGSIYSLPGGDFEGSEFVHAA 534
Db 414 GWQYAWKTEREDESQKE-----EGFPGRGSIYSLPGGDTGE-ADFVQSA 462
QY 535 LVQSALFSKGLABPRMSDAAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQLQFPA 594
Db 463 LVSPALYSKDLKKEHTIGPAMVHSE-TTKGSIWHDLPGVKRALVGVGLQLQFPA 521
QY 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLTLMPCIGFAMLLMDLSGR 654
Db 522 GINGVLYTTPQILEQAGVGILSNNGISSASASLISALTTFVMLPAIYAVAMRLMDLSGR 581
QY 655 RFLUGTIPILIASLVILVNSLIDGLTAHALLSTVSIVYFCCFVMGFGPIPNILCAE 714
Db 582 RTLLTIPILIASLVILVNSLIDGLTAHALLSTVSIVYFCCFVMGFGPAPNILCSE 641
QY 715 IFPTRVGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLKV 774
Db 642 IFPTRVGICIAICALTFWICDIIVTYSLPVLLKSLIGLAGVFGVYAVVCCISWVFFIKV 701
QY 775 PETKGMPELVITEFFAVGAKQAAA 798
Db 702 PETKGMPELVITEFFSVGARQAEA 725
```

```
RESULT 4
US-09-291-922-10
; Sequence 10, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
```

```
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-10
```

```
Query Match      41.0%; Score 1670.5; DB 4; Length 486;
Best Local Similarity 68.1%; Pred. No. 1.4e-156;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;
```

```
QY 308 DKEQITLYGPEEGSWIARPSKPIMLGSVLSLRHGS-MVNQSVPLMDPIVTLFGSVHE 367
Db 5 EKDQIKLYGPEEGQSWARPVAGPNSVG---LVSRKGSMAKNS-SLVDPVTLFGSVHE 59
QY 368 NMPQAGSMRSTLFPNFGSMFSVTQDHAKNEQWDEENLHRDDEEYASDAGGYEDNLHS 427
Db 60 KLPETG---STLFPHFGSMFSGVGNQPRNEDWDEESLAREGDDYVSD--AGDSDDLQOS 113
QY 428 PLISRQATGAEGKDIVHGHRSALSMRQTL-GEAGDVSTDTIGGQWLAWKSEKE 486
Db 114 PLISRQTTSLD-KDIPHAHNSLA-SMRQGLHGNSEPTGTGIGGQWLAWKSEKE 171
QY 487 GENGKKEGFKRYVLHQEGVPSRRGSIYSLPGGDFEGSEFVHAAALVQSALFSKGL 546
Db 172 GPDKKEGFKRYVLHQDGGSGSRGSIYSLP-CGDLPTDSEVVOAALVSPALYNEDL 230
QY 547 AEPMSDAAMVHPSEVAAGSRWKDLFEPGVRALLVGVGIQLQFAGINGVLYTTPQI 606
Db 231 MRQRPVGPAMVHPSETLAKGPSWDLFEPGVKHALIVGVMIQLQFSGINGVLYTTPQI 290
QY 607 LEQAGVAVILSKFGLSSASASILISLTLTLMPCIGFAMLLMDLSGRFLLGTIPILI 666
Db 291 LEQAGVYLLSLGLGSTSSFLISAVTTLMLPCIAIAMRLMDISGRRTLLSTIIPVLI 350
QY 667 ASLVILVNSLIDGLTAHALLSTVSIVYFCCFVMGFGPIPNILCAEIFPTRVGLCIA 726
Db 351 AALLILVLSLVLDGSIYANASISTISIVYFCEFWGFGPIPNILCAEIFPTRVGLCIA 410
QY 727 ICAFTWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLKVPETKGMPELVIT 786
Db 411 ICAFTWICDIIVTYSLPVMLNSVGLAGVFIYAVVCFIAWVFLKVPETKGMPELVIT 470
QY 787 EFPVAGAKQ 795
Db 471 EFPVAGAKQ 479
```

```
RESULT 5
US-09-291-922-14
; Sequence 14, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
```


	Best Local Similarity	21.9%	Pred. No. 4.7e-40;	Matches 165; Conservative 109; Mismatches 189; Indels 290; Gaps 18
QY	59	AVLVAIVASIGNLLQGWDNATIAA	AVLYIKKEFQJONE--PVEGLIVSM	SLIGATIVTT 116
Db	31	AFACAMLASMTSILG	YDIGVMSGAIIYIKRDLKVSDEQ	IEILGIINLYSLIGSCL--- 87
QY	117	ESGPLSDSIGRRPML	ILSSILYFESGLIMLSPNVYLLAR	FVDFGIGLAVTLVPLYI 176
Db	88	-AGRTSDWIGPRYTI	VFAGITFVVGALLMGFSPNYSFLM	FGFVAGIGIYALMIAPVYT 146
QY	177	SEIAPSEIRGLNLT	LPQFSGSGMFLSYCMVEGMS-LSP	SPDWIRMLGVLAIPSLFFFG 235
Db	147	AEVSPASSRGFLTS	FPEVFINGILLIGYISNYAFSK	ITLVKVMRMVLGVGAIPSV-L 205
QY	236	TIFFLPESPRWL	VSCKGRMAEAKVYLQKLRKDV	SGELSLLEGLVEVGGDTSEI 295
Db	206	GVLAMPESPRWL	VMRGRLGEARKVLNK-----	----- 232
QY	296	ATEAADLVTDG	KEQITLYGPEGQSWIARPSK	PIMLGSVISLASRHGSMVNQSV 355
Db	233	-----TSDSKEE	-----AQLRLA	----- 245
QY	356	DPIVTLFGSVHEN	MPQAGSGMSTLPNFGSMFS	VTDOHAKNEQWDEENLHRDDE 415
Db	246	-----EIKQAGIPESC	-----NDD	----- 260
QY	416	GAGGDYEDNLH	SPLSRQATGAEGKDIVHGH	RGSAISMRRQTLIGEGGDV 475
Db	261	-----	-----VVQVNKQS	-----NGEGV----- 273
QY	476	WQIAWKSEKEGEN	GRKEGFRVYLHQEGVPS	RGRGISVSLPGGDVFESEFV 535
Db	274	-----	-----	----- 273
QY	536	VSQSALFSKGLA	EPRMSDAAMVHPSEVA	AKSRWKDLF---EPGVRALL 592
Db	274	-----	-----WKELFLYFP	PAIRHIVIALGIIHFQO 300
QY	593	FAGINGVLYT	PQILEQAGVAVILSKFGL	SSASASILISLTLMLPCIG 652
Db	301	ASGVDAVVL	SPRIFEKAGITNDTHKL---	LATVAVGFVKIVFILA--- 351
QY	653	GRRFLLG	TIPIIASLVILVVS-NLID	LG-TLAHALSTVS-VIVY 708
Db	352	GRPLLSV	GVGMLSLTLTIAISLTV	IDHSERKLMWAVGSSIAM 411
QY	709	NILCAEIF	PTRVRLCIAICAFTE	WIGDIIIVTYSLPVMLNA 768
Db	412	WYSSSEIF	PLRLRAQGAAGAVAN	RTTSAVVSMTFLSLTRA 471
QY	769	FVFLKVPET	KMPLEVIT-EFFAV	GAQAQA 800
Db	472	FFYTVLP	PETRGKTLEDMEGS	FGTFPRSKSNASKA 504
RESULT 9				
US-09-291-922-20				
; Sequence 20, Application US/09291922				
; Patent No. 6383776				
; GENERAL INFORMATION:				
; APPLICANT: Allen, Steve				
; APPLICANT: Hiltz, Bill				
; APPLICANT: Kinney, Tony				
; APPLICANT: Tingey, Scott				
; TITLE OF INVENTION: Plant Sugar Transport Proteins				
; FILE REFERENCE: BB-1163				
; CURRENT APPLICATION NUMBER: US/09/291, 922				
; CURRENT FILING DATE: 1999-04-14				
; EARLIER APPLICATION NUMBER: 60/083, 044				
; EARLIER FILING DATE: April 24, 1998				
; NUMBER OF SEQ ID NOS: 30				
; SOFTWARE: Microsoft Office 97				

Db 339 VAVGVKTLFIL---VATFOLDKYGRRPLLTSVSGMIIAILTLAMSLTVIDHSHKIT 394
QY 683 LAHALLSTSVIVYFCCFVMGFRPIPNILCAEIFPTRVR--GLCIAICFTWIGDIIVT 740
Db 395 WAIALCITM-VCAVVASFSIGLGPITWVYSSEVFPLRLRAQGTSMGV-AVNRVSGVISI 452
QY 741 YSLPVMNAIGLAGVFSIYAVVCLISFVFVFLKVPETKMPLEVITEFF 789
Db 453 FFLP-LSHKITTGAFPLFGGIAIITAWFFLTLPETRGRTEENHELF 500

RESULT 12
US-09-291-922-22
; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22

Query Match 11.9%; Score 485; DB 4; Length 510;
Best Local Similarity 22.8%; Pred. No. 5.5e-39;
Matches 173; Conservative 98; Mismatches 187; Indels 300; Gaps 19;
QY 59 AVLVAIVASIGNLLOGWDNATIAAVALYIKKEFQLQNEPTVE---GLIVSMSLGATIVT 115
Db 21 AFACAILASMTSILGIDIGVSGASLIYIKDFNI-SDGKVEVLMGILNLVSLIGSFA-- 77
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVLY 175
Db 78 --AGRTSDWIGRRYTIFFAIVFFAGXFLMGFAVNYAMLNMGFRFVAGIGVYALMTAPVY 135
QY 176 ISEIAPSEIRGLNLTLPQFSGGMLSYCMVFGMSLP-SPDWIRMLGVLAIPSLFFFG 234
Db 136 TAEVSPASARGFLTSFPEVFINFGILLGYVSNYAFSRLPLNLGWRIMLGIGAAPSV-LLA 194
QY 235 LTIFFLPESPRWLVSKGMAEAKVLOKLRGKDDVSGELSLLEGLVGGDTSEIYYIIG 294
Db 195 LMTLGMPESPRWLVMMKGRADAKVLEKT-----SDT----- 226
QY 295 PATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASRHGSMVNVQSVFL 354
Db 227 -AEAAERL----- 234
QY 355 MDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFSVTDOHAKNEQWDEENLHRDEEYAS 414
Db 235 ----- 234
QY 415 DGAGDYEDNLHPLSLRQATGAEKDIVHHGHRGALSMMRQOTLLGEGGDVGSSTDIG 474
Db 235 -----ADTKA 239
QY 475 GWQIAWKSEKEGENGRKEGFKRVIYLHQEGVPGSRRGSIVSLP--GGGDVFEGSEFVHA 532
Db 240 A-----AGIPEELDGDVVTVPKRGSGN----- 261

QY 533 AALVSQALFSKGLAEPRMDSAMVHPSEVAAKGRMKDLF---EPGVRRLVGVGICQI 589
Db 262 -----EKRV-----WKELILSPTAMRRILLSGIGIHF 289
QY 590 LOQFAGINGVLYTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGPAMLLM 649
Db 290 FOHALGHSVVFSPLVFKSPGLT-NDKHFLGTWPFGVTKRLFILL-----ATFFI 340
QY 650 DLSGRFLLGTIPILASLV-----ILVSNLIDGLTAHALLSTSVIVYFCCFVMGF 704
Db 341 DGVRRPRLLGSTGIIISLIGLAGLTVVGQHPDAKI PWAIGLSIASTLAVAFFSIGL 400
QY 705 GPIPNILCAEIFPTRVR--GLCIAICFTWIGDIIVTYSLPVMLNAGLAGVFSIYAVV 762
Db 401 GPITWVYSSEIFPLQVRALGCSLGVAANRVTSGVISMTFL--SLSKAITIGGSFFLYSGI 458
QY 763 CLISFVFVFLKVPETKMPLEVITEFFAVGAKQAATAKA 800
Db 459 AALAWVFYTYLTPETRGRTEEMSKLF--GDTAAASES 494

RESULT 13
US-09-291-922-26
; Sequence 26, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-26

Query Match 11.4%; Score 463.5; DB 4; Length 539;
Best Local Similarity 21.7%; Pred. No. 8.2e-37;
Matches 166; Conservative 108; Mismatches 200; Indels 291; Gaps 17;
QY 52 LEDKMSG---AVLVAIVASIGNLLOGWDNATIAAVALYIKKEFQLQNEPTVE---GLIV 104
Db 35 VEPKKKKGNRFACAILASMTSILGIDIGVSGASLIYIQDKLI-NDTQLEVLMGILN 93
QY 105 SMSLICATIVTTFSGPLSDSIGRRPMLILSSILYFSGLIMLWSPNVYVLLARFVDGFG 164
Db 94 VYSLIGSFA---AGRTSDWIGRRFTIVFAAVIFFAGALIMGFSVNYAMLNMGFRFVAGIG 149
QY 165 IGLAVTLVPLYISEIAPSEIRGLNLTLPQFSGGMLSYCMVFGMS-LSPPDWIRIMLG 223
Db 150 VGVALMTAPVNTGEVSPASARGVLTSPPEVFINFGILLGYVSNFAFARLSTRLGWRIMLG 209
QY 224 VLAIPSLFFFGTLTIFFLPESPRWLVSKGMAEAKVLOKLRGKDDVSGELSLLEGLVVG 283
Db 210 IGAVPSV-LIAFMVLGMPESPRWLVMMKGRADAKVLAKT----- 248
QY 284 GDTSEIYYIIGPATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASR 343
Db 249 SDT-----PEEAERIA----- 260
QY 344 HGSMTNVQSVPLMDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFSVTDOHAKNEQWDEE 403
Db 261 ----- 260

QY 404 NLRDDEEYASDAGGDYEDNLHSPLLSRQATGAEGKDIYHGHRSALSMRQTLLEG 463
Db 261 ----DICTAA----- 266
QY 464 GDGVSTDIGGQWLAKWKSEKENGKRGKGFYVYHOGVPGSRGSIYSLPGGDIV 523
Db 267 -----GIPGLDGDVVPVK---- 281
QY 524 FEGSEFVHAALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRRALLV 583
Db 282 -----NKGSEKRVKDLILSPITIA-----MRHILIA 309
QY 584 GVGIOILQOFAGINGVLYYTPQILEQAGVAVILSKFGLS-ASASILISLTLMLPCT 642
Db 310 GIGIHFFQSSGIDAVVLYSPVFKSAGIT-----GDSRLRGTTVAVGATNTVFIL--- 360
QY 643 GFAMLMDLSGRRFLLGITIPILIASLV-----ILVVSNLIDLGLAHLASTVSIVYF 697
Db 361 -VATFLLDRIRRRPLVLTSTGMLVSLVGLATGLTVISRHPDEKITWAIVLCFICIMAYV 419
QY 698 CCFVWGFGPIPNILCAEIPFTRVGLCIAICFTFNIGDIIVTYSLPVMLNAIGLAVES 757
Db 420 AFSIGLGPITWVYSSEIFPLHVALGCSLGAVANRLTSGVISMFTFISLSKAMTIGAF 479
QY 758 IYAVVCLISFVFLKVPETKMPLEVITEFFAVGA--KQAAKA 800
Db 480 LFAGIASFAWVFAYLPLETRGRLEDMSSLFGNTATHKQGAEA 524

RESULT 14
US-08-928-692-13
; Sequence 13, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59587270 No. 5958727disk of No. 5958727th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEPT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5958727e
; US-08-928-692-13

Query Match 10.3%; Score 421.5; DB 2; Length 584;
Best Local Similarity 19.1%; Pred. No. 1.4e-32;
Matches 150; Conservative 121; Mismatches 195; Indels 321; Gaps 19;
QY 44 EPTVSD-----IIEDKMSGAVL-VAIVASIGNLLQGWDNATIAAVALYIKKEFOLONE 96
Db 64 KPVNEDDTSVMITFNQSLSPFIITLTFVASISGFMEGYDTGYISSALISIGTD--LDHK 121
QY 97 PVEG--LIVSMSLIGATIVTTFSGPLSDSIGRRPMLILSLYFFSGILMLSPNVY 153
Db 122 VLTGKEKIVTAATSGALITSIFAGTAADIPGRKCLMGSNLMFVIGAILQVSAHTFWQ 181
QY 154 LLARFVDFGIGLAVTLVPLYISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLS 213
Db 182 MAVGRLLMGFGVIGISLAPLFISEIAPKMRGLTVINSLMLTGGQLVAYCGAGLNY- 240
QY 214 PSPDWRIMGLVLAIPSLFFEGLTIFYLPESPRWLVSGRMAEAKKVLQKRGKDDVSGEL 273
Db 241 VNNGWRLVGLSLIPTAVQF-TCLCFLPDTPRYVMKGDLARATEVLK--RSYTDISEI 297
QY 274 SLLEGLEVGSDTSEIIEYIIGPATEADDLVTDGDKQITLYGPEEGSNIAKPSKPIIM 333
Db 298 -----IERKV-----BELVT----- 307
QY 334 LGSVLSLASRHGSMVNOQVPLMDPIVTLFGSVHNNPQAGGSMRSTLPNFGSMFSVTQ 393
Db 308 -----LNQSI----- 313
QY 394 HAKNEQWDEENLHRDDEEYASDAGGDYEDNLHSPLLSRQATGAEGKDIYHGHRSALS 453
Db 314 -----GKNV----- 317
QY 454 MRQTLLEGGDVSTDIGGQWLAKWKSEKENGKRGKGFYVYHOGVPGSRGS 513
Db 318 ----- 317
QY 514 IVSLPGGDVFESEFVHAALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLF 573
Db 318 -----PEKV-----WNTIK 326
QY 574 E---PGVRALLVGVGIQILQOFAGINGVLYTPQILEQAGVAVILSKFGLSSASASIL 629
Db 327 ELHTVPSNLRALIGGLQAIQOFTGWNLSMFGSTIFETVGFK-----NSSAVSII 378
QY 630 ISSLTLLMLPCIGFAMLMDLSGRRFLLGITIPILIASLVL-----VVSN 676
Db 379 VSGTNEIFTL--VAFFSIDKIGRTILLGPGMTMALVVCISIAFHGLIKPDGAVAV 434
QY 677 LIDLGLAHLASTVSIVVYFCFVWGFGPIPNILCAEIPFTRVGLCIAICFTFWIGD 736
Db 435 VSSGFSWGIIVIFIVFAFALGIGTVP-WQSELPQNVRGIGTSVATATNWAGS 493
QY 737 IIVTYSLPVMLNAIGLAVESIVAVVCLISFVFLKVPETKMPLE---VITEFFAVG 792
Db 494 LVIASTFLTMQNTITPAGTFAFAGLSCLSTIFCYFCYBELSGLELEEVQTLKDGFNK 553
QY 793 AKQAAK 799
Db 554 ASKALAK 560

RESULT 15
US-09-339-972-13
; Sequence 13, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80

CORRESPONDENCE ADDRESS:
ADDRESS: No. 6323002o No. 6323002disk of No. 6323002th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/339,972
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/928,692
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4944.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6323002e
US-09-339-972-13

Query Match 10.3%; Score 421.5; DB 4; length 584;
Best Local Similarity 19.1%; Pred. No. 1.4e-32;
Matches 150; Conservative 121; Mismatches 195; Indels 321; Gaps 19;

QY 44 ERYTSD-----ILEDKMSGAVL-VAIVASIGNLQGMNATIAAVALYIKKEFOLONE 96
Db 64 KPVNDEDDTSMITFNQSLSPITITLTFVASISGFMGYDTGYISSALISIGTD--LDHK 121
QY 97 PVEG---LIVSMILGATVITTFSGPLSDSIGRRPMLISSILYFSGLIMLWSPNYV 153
Db 122 VLTGEKEIVTAATSLGALITSIFAGTAADIFGRKRCMLGSLNMFVIGAILQVSAHTFWQ 181
QY 154 LLARFVDGFIGLAVTLVPLYSEIAPSEIRGLNLTLPQSGSGMFLSYCMVFGMSLS 213
Db 182 MAVGRLIMFGVIGSLIAPLFISEIAPKMRGLTVINSIMLTGQLVAYGCGALNY- 240
QY 214 PSPDWRIMGLVLAIPSLFFGLTIFYLPESPRWLVSKGMAEAKVLOKLRGKDVSGEL 273
Db 241 VNNGWRLILVGLSLIPTAVQF-FCUCFLPDTPRYYVMKGDILARATEVLK--RSYTDTSSEI 297
QY 274 SLLLEGLEVGDTSEIEYIIGPATEAADLVTDGDKEQITLYGPBEGQSWIARPSKPIIM 333
Db 298 -----IERKV-----ELVT----- 307
QY 334 LGSVLSLSRHSVMNOSVPLMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQ 393
Db 308 -----LNQSI----- 313
QY 394 HAKNEQWDEENLHRDEEYASDAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHSALS 453
Db 314 -----GKNV----- 317
QY 454 MRROTLLGEGGDVSTDIGGQWLAMKWKSEKGENGRKEGFKRVYLHQEGVPSGRGS 513
Db 318 ----- 317
QY 514 IVSLPGGDVFEQSEFVHAALVQSALFSKGLAEPMSDAAVHPSEVAAGSRWKDLF 573

Db 318 -----PERV-----WNTIK 326
QY 574 E----PGVRRLIVGVIQILOQFAGINGVLYTPQILEQAGVAVILSKGLSSASISL 629
Db 327 ELHTVPSNLRALIIGCGLQAIQOFTGWSNLVYFSGTIFETVGFK-----NSSAVSI 378
QY 630 ISSLTLLMLPCIGFAMLLMDLSGRRFLIGTIPILLASLVIL-----VWSN 676
Db 379 VSGTNFIFTL---VAFFSIDKIGRTILLIGLPMTMALVVCIAHFHIGIKFDGAVAV 434
QY 677 LIDLGLAHALLSTVSVIVYFCCFVMGFGPIENILCAEIFPTRVRGLCIACAFTWIGD 736
Db 435 VSSGFSWGIIVIFIVFAFALGIGYF-WQOSELPQNVRGIGTSYATATNWAGS 493
QY 737 IIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLKVPETKMPLE---VITEFFAVG 792
Db 494 LVIASTFLTMLQNTITPAGTFAPFAGLSCLSTIFCYFCYPELSGLELEEVQTIKDFNIK 553
QY 793 AKQAAK 799
Db 554 ASKALAK 560

Search completed: January 5, 2004, 18:59:28
Job time : 37 secs

THIS PAGE BLANK (05710)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 14:13:32 ; Search time 867 Seconds
(without alignments)
11086.463 Million cell updates/sec

Title: US-10-051-909-31

Perfect score: 2777
Sequence: 1 cacggggttagattcggagc.....aaaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2777	100.0	2777	14 US-10-051-909-31	Sequence 31, Appl
2	2740.4	98.7	2908	12 US-10-051-902-1	Sequence 350, App
3	2200.2	79.2	2824	14 US-10-051-902-1	Sequence 1, Appli
4	2200.2	79.2	2824	14 US-10-051-902-1	Sequence 1, Appli
5	940.2	33.9	2601	14 US-10-051-902-7	Sequence 7, Appli
6	940.2	33.9	2601	14 US-10-051-902-7	Sequence 7, Appli
7	875.8	31.5	2190	10 US-09-938-842A-1315	Sequence 1315, Ap
8	700	25.2	2205	10 US-09-938-842A-2254	Sequence 2254, Ap
9	691.2	24.9	1487	14 US-10-051-902-13	Sequence 13, Appl
10	691.2	24.9	1487	14 US-10-051-902-13	Sequence 13, Appl
11	594.6	21.4	1692	14 US-10-051-902-9	Sequence 9, Appli
12	594.6	21.4	1692	14 US-10-051-902-9	Sequence 9, Appli
13	394.4	14.2	1009	14 US-10-051-902-15	Sequence 15, Appl
14	394.4	14.2	1009	14 US-10-051-902-15	Sequence 15, Appl
15	269.2	9.7	285	9 US-09-923-876-6145	Sequence 6145, Ap

16	269.2	9.7	285	12 US-09-923-876-6145	Sequence 6145, Ap
17	257.2	9.3	282	9 US-09-294-093B-699	Sequence 699, App
18	227	8.2	870	14 US-10-051-902-5	Sequence 5, Appli
19	227	8.2	870	14 US-10-051-902-5	Sequence 5, Appli
20	123.2	4.4	510	14 US-10-051-902-11	Sequence 11, Appl
21	123.2	4.4	510	14 US-10-051-902-11	Sequence 11, Appl
22	114.2	4.1	443	14 US-10-051-902-3	Sequence 3, Appli
23	114.2	4.1	443	14 US-10-051-902-3	Sequence 3, Appli
24	112.4	4.0	299	9 US-09-294-093B-6094	Sequence 6094, Ap
25	96.2	3.5	1853	14 US-10-051-902-23	Sequence 23, Appl
26	96.2	3.5	1853	14 US-10-051-902-23	Sequence 23, Appl
27	84	3.0	2017	14 US-10-051-902-21	Sequence 21, Appl
28	84	3.0	2017	14 US-10-051-902-21	Sequence 21, Appl
29	82.6	3.0	2051	12 US-10-051-154-356	Sequence 356, App
30	81.4	2.9	260	10 US-09-878-574-9951	Sequence 9951, Ap
31	78.2	2.8	2089	14 US-10-051-902-25	Sequence 25, Appl
32	78.2	2.8	2089	14 US-10-051-902-25	Sequence 25, Appl
33	77.6	2.8	1482	10 US-09-938-842A-795	Sequence 795, App
34	75	2.7	1644	10 US-09-938-842A-320	Sequence 320, App
35	66.2	2.4	1386	12 US-10-369-493-47058	Sequence 47058, A
36	65.8	2.4	1872	14 US-10-051-902-27	Sequence 27, Appl
37	65.8	2.4	1872	14 US-10-051-902-27	Sequence 27, Appl
38	65.2	2.3	1655	12 US-10-369-493-26996	Sequence 26996, A
39	64.2	2.3	1422	12 US-10-369-493-46784	Sequence 46784, A
40	63.4	2.3	1377	12 US-10-369-493-39348	Sequence 39348, A
41	63.4	2.3	1377	12 US-10-369-493-39348	Sequence 39348, A
42	63.4	2.3	1419	12 US-10-369-493-38982	Sequence 38982, A
43	63.4	2.3	1914	14 US-10-051-902-19	Sequence 19, Appl
44	63.4	2.3	1914	14 US-10-051-902-19	Sequence 19, Appl
45	63.4	2.3	2127	12 US-10-310-154-354	Sequence 354, App

ALIGNMENTS

Handwritten signature

RESULT 1
US-10-051-909-31
Sequence 31, Application US/10051909
Publication No. US20020199217A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: B1163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 31
LENGTH: 2777
TYPE: DNA
ORGANISM: Zea mays
US-10-051-909-31

QY	1	CACGGGGTTAGATTCCGAGCGGCTCTTGCTTGACGTCACCAAGCCCTTACCCCTGATC	60
DB	1	CACGGGGTTAGATTCCGAGCGGCTCTTGCTTGACGTCACCAAGCCCTTACCCCTGATC	60
QY	61	TGACCCGAGGAGCGGCTCTTCCGTCAGTTGTTCTTGCTTGCGCTGCGCTCTTCCGC	120
DB	61	TGACCCGAGGAGCGGCTCTTCCGTCAGTTGTTCTTGCTTGCGCTGCGCTCTTCCGC	120
QY	121	CTGCTTCGTGTTCTTCAAGGAGCCGGTGACCTCGACGATATCTTGAGGACAGATGT	180

Db 121 CTGCTTCGTCTTCTTACAGAGACCGGTGACCTCGACGATATCTTGGAGACAAGATGT 180
QY 181 CGGGGGCTGTCTTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGCAAGGGGTGGACA 240
Db 181 CGGGGGCTGTCTTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGCAAGGGGTGGACA 240
QY 241 ATGCCACATCGCAGCTGCTGTCTGTATATAAGAAAGGAATTTCAATTGCCAAAATGAGC 300
Db 241 ATGCCACATCGCAGCTGCTGTCTGTATATAAGAAAGGAATTTCAATTGCCAAAATGAGC 300
QY 301 CCACGTGAGAGGGAATAATTGTCAATGTCACTTATCGGCGCCACCATCGTTACTACAT 360
Db 301 CCACGTGAGAGGGAATAATTGTGTCAATGTCACTTATCGGCGCCACCATCGTTACTACAT 360
QY 361 TCTCCGGGCAATTATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTTCAATTTC 420
Db 361 TCTCCGGGCAATTATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTTCAATTTC 420
QY 421 TGTACTTCTTACGGGCTCATCATGCTATGGTCTCTTAATGTCTAATGCTGCTGTGG 480
Db 421 TGTACTTCTTACGGGCTCATCATGCTATGGTCTCTTAATGTCTAATGCTGCTGTGG 480
QY 481 CACGCTTCGTAGATGGATTGGTATGGCTTGGCTGTCACGCTGTGCTTTGTACATTT 540
Db 481 CACGCTTCGTAGATGGATTGGTATGGCTTGGCTGTCACGCTGTGCTTTGTACATTT 540
QY 541 CAGAAATAGCCCCCTTCGAGATTAGAGTTGCTGAATACACTACACAATTCAGTGGAT 600
Db 541 CAGAAATAGCCCCCTTCGAGATTAGAGTTGCTGAATACACTACACAATTCAGTGGAT 600
QY 601 CAGAGGAATGTTCTTGTCACTGCAATGGTGTGATGTCCTGTGCGCCATCACCCG 660
Db 601 CAGAGGAATGTTCTTGTCACTGCAATGGTGTGATGTCCTGTGCGCCATCACCCG 660
QY 661 ATTGAGAAATATATGCTGTGTCGCGATACCTTCATTGTTCTTCTTGGTTGACAA 720
Db 661 ATTGAGAAATATATGCTGTGTCGCGATACCTTCATTGTTCTTCTTGGTTGACAA 720
QY 721 TATTTTATCTTCTTGAATCTCCAAAGATGGCTCGTTAGCAAAAGTGGATGGCAGAGGCAA 780
Db 721 TATTTTATCTTCTTGAATCTCCAAAGATGGCTCGTTAGCAAAAGTGGATGGCAGAGGCAA 780
QY 781 AAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATGTCCCTTCTTC 840
Db 781 AAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATGTCCCTTCTTC 840
QY 841 TCGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGATACATTTGACCTGCCA 900
Db 841 TCGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGATACATTTGACCTGCCA 900
QY 901 CCGAGGACCGCATGATCTTGTATGACGGTGATTAAGAAACAATCACACTTATGGGC 960
Db 901 CCGAGGACCGCATGATCTTGTATGACGGTGATTAAGAAACAATCACACTTATGGGC 960
QY 961 CTGAAGAAAGCCAGTCATGATTGCTCGACCTTCTAAGGGAACCATCATGCTTGAAGTG 1020
Db 961 CTGAAGAAAGCCAGTCATGATTGCTCGACCTTCTAAGGGAACCATCATGCTTGAAGTG 1020
QY 1021 TGTCTTCTTCTTGCATCTCGTCATGGAAGCATGTTGACCAAGAGTGAACCCCTTATGATC 1080
Db 1021 TGTCTTCTTCTTGCATCTCGTCATGGAAGCATGTTGACCAAGAGTGAACCCCTTATGATC 1080
QY 1081 CGATTGTGACACTTTTGTAGTGTCCATGAGAATATGCTCAAGCTGGAGGAAGTATGA 1140
Db 1081 CGATTGTGACACTTTTGTAGTGTCCATGAGAATATGCTCAAGCTGGAGGAAGTATGA 1140
QY 1141 GGAGCAATGTTTCCAAACTTTGAGATATGTTCAAGTGTCAAGATCAGATGCCAAA 1200
Db 1141 GGAGCAATGTTTCCAAACTTTGAGATATGTTCAAGTGTCAAGATCAGATGCCAAA 1200
QY 1201 ATGAGCAATGGGATGAAGAAATCTTCAATAGGATGACGAGAGTAAGCATCTGATGTTG 1260
Db 1201 ATGAGCAATGGGATGAAGAAATCTTCAATAGGATGACGAGAGTAAGCATCTGATGTTG 1260

QY 1261 CAGAGGTGACTATGAGCAAAATCTCCATAGCCCATGTGCTCAGGCAAGCAACAGGTG 1320
Db 1261 CAGAGGTGACTATGAGCAAAATCTCCATAGCCCATGTGCTCAGGCAAGCAACAGGTG 1320
QY 1321 CGGAAGGAAGACATTTGTGCAACCATGTGTACCCGTGGAAGTGTCTTGAAGATGAAGGC 1380
Db 1321 CGGAAGGAAGACATTTGTGCAACCATGTGTACCCGTGGAAGTGTCTTGAAGATGAAGGC 1380
QY 1381 AAACCTCTTAGGGAGGGTGAGATGTTGTAGACAGCACTGATATCGGTGGGGATGGC 1440
Db 1381 AAACCTCTTAGGGAGGGTGAGATGTTGTAGACAGCACTGATATCGGTGGGGATGGC 1440
QY 1441 AGCTTGCTTGGAAATGGTCAAGAAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGTTCA 1500
Db 1441 AGCTTGCTTGGAAATGGTCAAGAAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGTTCA 1500
QY 1501 AAAGAGTCTACTTGCACCAAGAGGAGTTCTGTGCTCAAGAAAGGGCTCAATTGTTTCA 1560
Db 1501 AAAGAGTCTACTTGCACCAAGAGGAGTTCTGTGCTCAAGAAAGGGCTCAATTGTTTCA 1560
QY 1561 TTCCCGGTGGCGCATGTTTGTAGGGTAGTGAATTTGTACATGCTGCTGCTTAACTAA 1620
Db 1561 TTCCCGGTGGCGCATGTTTGTAGGGTAGTGAATTTGTACATGCTGCTGCTTAACTAA 1620
QY 1621 GTCAGTACGACCTTTTCTCAAAAGGCTTGTCTGAACACGATGTCAAGTGTGCAATG 1680
Db 1621 GTCAGTACGACCTTTTCTCAAAAGGCTTGTCTGAACACGATGTCAAGTGTGCAATG 1680
QY 1681 TTCAACCCATCTGAGGTAGCTGCCAAAGGTTCAAGTTGGAAGAAAGTTGTTGAACCTGGAG 1740
Db 1681 TTCAACCCATCTGAGGTAGCTGCCAAAGGTTCAAGTTGGAAGAAAGTTGTTGAACCTGGAG 1740
QY 1741 TGAGGCGTGCCCTGTAGTGGGTGTGGAATTCAGATCTTCAACAGTTTGTCTGGAATTA 1800
Db 1741 TGAGGCGTGCCCTGTAGTGGGTGTGGAATTCAGATCTTCAACAGTTTGTCTGGAATTA 1800
QY 1801 ACGGTGTCTGTACTATAACCCACAATTTCTTGAGCAAGCTGTGTGGCAGTTATCTTT 1860
Db 1801 ACGGTGTCTGTACTATAACCCACAATTTCTTGAGCAAGCTGTGTGGCAGTTATCTTT 1860
QY 1861 CCAAAATTTGCTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACATCCTTAC 1920
Db 1861 CCAAAATTTGCTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACATCCTTAC 1920
QY 1921 TAAATGCTTCTTGCATGCTGCTTGGCATGCTGCTTAAAGATCTTCCGAAGAGGTTT 1980
Db 1921 TAAATGCTTCTTGCATGCTGCTTGGCATGCTGCTTAAAGATCTTCCGAAGAGGTTT 1980
QY 1981 TGTGCTAGGACAATTCATCTTGATAGCATCTCTAGTTATCTCTGTTGTCCAAATC 2040
Db 1981 TGTGCTAGGACAATTCATCTTGATAGCATCTCTAGTTATCTCTGTTGTCCAAATC 2040
QY 2041 TAAATGATTGGGTACACTAGCCCATGCTTGTCTTCCACCGTCAGTGTATCTCTACT 2100
Db 2041 TAAATGATTGGGTACACTAGCCCATGCTTGTCTTCCACCGTCAGTGTATCTCTACT 2100
QY 2101 TCTGCTGCTTGTATGGGATTTGGTCCCATCCCAACATTTATGTGACAGAAATCTTTC 2160
Db 2101 TCTGCTGCTTGTATGGGATTTGGTCCCATCCCAACATTTATGTGACAGAAATCTTTC 2160
QY 2161 CAACCAAGGTTCTGCGCTCTGTATTGCAATTTGTGCTTACATCTGATCGAGATA 2220
Db 2161 CAACCAAGGTTCTGCGCTCTGTATTGCAATTTGTGCTTACATCTGATCGAGATA 2220
QY 2221 TCAATGCTACCTACAGCCTTCTGTGATGCTGAATGCTATTGGACTGCGGGGTGTTTCA 2280
Db 2221 TCAATGCTACCTACAGCCTTCTGTGATGCTGAATGCTATTGGACTGCGGGGTGTTTCA 2280
QY 2281 GCATATATGCAAGTCGTATGCTTGAATTTCTTGTGTGTTCTTCTTAAAGTCCCTGAGA 2340
Db 2281 GCATATATGCAAGTCGTATGCTTGAATTTCTTGTGTGTTCTTCTTAAAGTCCCTGAGA 2340

Db 759 AGGAATGTTCTGTCTACTGTCATGTTGTTGGGATGTCCCTGTCCCATCACCAGATTG 818
QY 665 GAGAAATTAATGCTTGGTGTGCTCGCGAATACCTTCATGTGTTCTTTGGTTTGACAATATT 724
Db 819 GAGAAATTAATGCTTGGTGTGCTCGCGAATACCTTCATGTGTTCTTTGGTTTGACAATATT 878
QY 725 TTATCTTCTGAATCTCCAAAGATGGCTGTAGCAAAAGTCCGATGGCAGAGCAAAAAA 784
Db 879 TTATCTTCTGAATCTCCAAAGATGGCTGTAGCAAAAGTCCGATGGCAGAGCAAAAAA 938
QY 785 GGTGTGCAAAAGTTACGGGGAAAGCATGTCTCAGGTGAATTGTCCTTCTCTCGA 844
Db 939 GGTGTGCAAAAGTTACGGGGAAAGCATGTCTCAGGTGAATTGTCCTTCTCTCGA 998
QY 845 AGGGTTGAGGTTGAGAGAGACATTCCATTGAAGATACATCATTTGACCCTGCCACCGA 904
Db 999 AGGGTTGAGGTTGAGAGAGACATTCCATTGAAGATACATCATTTGACCCTGCCACCGA 1058
QY 905 GGCAGCCGATGATCTTGTACTGACGGTGAATAAGAACAAATCACACTTTATGGCCCTGA 964
Db 1059 GGCAGCCGATGATCTTGTACTGACGGTGAATAAGAACAAATCACACTTTATGGCCCTGA 1118
QY 965 AGAAGGCCAGTCAATGATGATGCTCGACCTTTCTAAGGGACCATCATGCTTGAAGTGTCT 1024
Db 1119 AGAAGGCCAGTCAATGATGATGCTCGACCTTTCTAAGGGACCATCATGCTTGAAGTGTCT 1178
QY 1025 TTTCTTTGATCTCGTCAATGGAGCATGGTGAACCAGAGTGTACCCCTTATGATCCGAT 1084
Db 1179 TTTCTTTGATCTCGTCAATGGAGCATGGTGAACCAGAGTGTACCCCTTATGATCCGAT 1238
QY 1085 TGTGACATTTTGTGTAGTGTCCATGAGAAATATGCTCAAGCTGAGGAAAGTATGAGAG 1144
Db 1239 TGTGACATTTTGTGTAGTGTCCATGAGAAATATGCTCAAGCTGAGGAAAGTATGAGAG 1298
QY 1145 CACATGTTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGATGCCAAATGA 1204
Db 1299 CACATGTTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGATGCCAAATGA 1358
QY 1205 GCAGTGGGATGAAGAAATCTTCATAGGAGTGAAGAGAGTACGCATCTGATGTCAGG 1264
Db 1359 GCAGTGGGATGAAGAAATCTTCATAGGAGTGAAGAGAGTACGCATCTGATGTCAGG 1418
QY 1265 AGGTGACTATGAGCAATCTCCATAGCCATTGCTGTCCAGGCAAGCAACAGGTGCGGA 1324
Db 1419 AGGTGACTATGAGCAATCTCCATAGCCATTGCTGTCCAGGCAAGCAACAGGTGCGGA 1478
QY 1325 AGGGAAGGACATGTGCAACCATGTCACCGTGAAGTGTGTCATGAGATGAAGGCAAAAC 1384
Db 1479 AGGGAAGGACATGTGCAACCATGTCACCGTGAAGTGTGTCATGAGATGAAGGCAAAAC 1538
QY 1385 CCTTTAGGGGAGGTTGAGATGGTGTGAGCAGCCTGATATCGGTGGGGGATGCGAGCT 1444
Db 1539 CCTTTAGGGGAGGTTGAGATGGTGTGAGCAGCCTGATATCGGTGGGGGATGCGAGCT 1598
QY 1445 TGCCTTGAATAATGTCAGAGAAAGGTTGAAATGTTAGAAAGGAGTGTTCAAAAG 1504
Db 1599 TGCCTTGAATAATGTCAGAGAAAGGTTGAAATGTTAGAAAGGAGTGTTCAAAAG 1658
QY 1505 AGTCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAAAGGGGCTCAATTGTTCACTTCC 1564
Db 1659 AGTCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAAAGGGGCTCAATTGTTCACTTCC 1718
QY 1565 CGGTGTGCGCATGTTTTTGAAGGTTAGTGAATTTGACATGCTGCTTTAGTAAGTCA 1624
Db 1719 CGGTGTGCGCATGTTTTTGAAGGTTAGTGAATTTGACATGCTGCTTTAGTAAGTCA 1778
QY 1625 GTGAGCACTTTTCTCAAGAGGCTTGTGTAACACGACATGTCAATGCTGCCATGTTCA 1684
Db 1779 GTGAGCACTTTTCTCAAGAGGCTTGTGTAACACGACATGTCAATGCTGCCATGTTCA 1838
QY 1685 CCCATCTAGAGTAGCTGCCAAAGGTTCAAGTTGGAAGAATTTGTTGAACCTGAGGTGAG 1744
Db 1839 CCCATCTAGAGTAGCTGCCAAAGGTTCAAGTTGGAAGAATTTGTTGAACCTGAGGTGAG 1898

QY 1745 GCGTCCCTGTTAGTCGGTGTGGAATTAGATCCTTCAACAGTTTGCTGGAATAAACGG 1804
Db 1899 GCGTCCCTGTTAGTCGGTGTGGAATTAGATCCTTCAACAGTTTGCTGGAATAAACGG 1958
QY 1805 TGTCTGTACTATACCCCAAAATTTCTGAGCAAGCTGTGTGGCAGTTATCTTTCCAA 1864
Db 1959 TGTCTGTACTATACCCCAAAATTTCTGAGCAAGCTGTGTGGCAGTTATCTTTCCAA 2018
QY 1865 ATTTGCTCTAGCTCGGCATCAGCATCCATCTTGAATCAGTTCTCTACTACTTAATAAT 1924
Db 2019 ATTTGCTCTAGCTCGGCATCAGCATCCATCTTGAATCAGTTCTCTACTACTTAATAAT 2078
QY 1925 GCTTCTTGCATTTGGCTTTGCCATGCTGCTTATGATCTTTCCGAAAGAGTTTTCCT 1984
Db 2079 GCTTCTTGCATTTGGCTTTGCCATGCTGCTTATGATCTTTCCGAAAGAGTTTTCCT 2138
QY 1985 GCTAGGCACAATTCATCTTGAATAGCATCTTATGTTATCTGTTGTCTCAATCTAAT 2044
Db 2139 GCTAGGCACAATTCATCTTGAATAGCATCTTATGTTATCTGTTGTCTCAATCTAAT 2198
QY 2045 TGAATTTGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAAGTGTATCTGTACTCTG 2104
Db 2199 TGAATTTGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAAGTGTATCTGTACTCTG 2258
QY 2105 CTGCTTCTGTTATGGGATTTGGTCCCATCCCAACAATTTATGTGACAGATCTTTCCAAC 2164
Db 2259 CTGCTTCTGTTATGGGATTTGGTCCCATCCCAACAATTTATGTGACAGATCTTTCCAAC 2318
QY 2165 CAGGGTTCGTGGCCCTGTGATTTGCCATTTGTGCTTTTACATTTCTGGATCCGAGATATCAT 2224
Db 2319 CAGGGTTCGTGGCCCTGTGATTTGCCATTTGTGCTTTTACATTTCTGGATCCGAGATATCAT 2378
QY 2225 GCTCACCTACAGCCTTCTGTGATGCTGAAATGCTAATGGAAGTGGCGGCTTTTACAGCAT 2284
Db 2379 GCTCACCTACAGCCTTCTGTGATGCTGAAATGCTAATGGAAGTGGCGGCTTTTACAGCAT 2438
QY 2285 ATATGAGTGTATGCTGATTTCTTGTGTGCTTCTTAAAGTCCCTGAGCAAAA 2344
Db 2439 ATATGAGTGTATGCTGATTTCTTGTGTGCTTCTTAAAGTCCCTGAGCAAAA 2498
QY 2345 GGGGATGCCCCCTGAGGTTATTAACCAATTTCTTGCAAGTTGGTGCAGAACGCGCTGC 2404
Db 2499 GGGGATGCCCCCTGAGGTTATTAACCAATTTCTTGCAAGTTGGTGCAGAACGCGCTGC 2558
QY 2405 AAAAGCCTAATTTCTTTGGTACCTTTGTGTGCAACTATTGCACTGTAAGTTAGAACTTG 2464
Db 2559 AAAAGCCTAATTTCTTTGGTACCTTTGTGTGCAACTATTGCACTGTAAGTTAGAACTTG 2618
QY 2465 AAGGGGTTTCAACCAAGAGCTCGGAGAAATTAATTGGAATTTGTGTAATGTTAAGGAAAC 2524
Db 2619 AAGGGGTTTCAACCAAGAGCTCGGAGAAATTAATTGGAATTTGTGTAATGTTAAGGAAAC 2678
QY 2525 GAACATCTGCTCATGCTCTCAAAACGGTAAAAAAGAGTCCCTCAATGGCAATAGAGTGC 2584
Db 2679 GAACATCTGCTCATGCTCTCAAAACGGTAAAAAAGAGTCCCTCAATGGCAATAGAGTGC 2738
QY 2585 GTTAAGTTGTCAATGCTATTAACCATATGTTTACCTAATTTGTACTGATATTAAGTCAA 2644
Db 2739 GTTAAGTTGTCAATGCTATTAACCATATGTTTACCTAATTTGTACTGATATTAAGTCAA 2798
QY 2645 GCTATTCAACGCTGTTGTGCTAGAAATCTTTAGAACAAAGATGATATGATCTGATCT 2704
Db 2799 GCTATTCAACGCTGTTGTGCTAGAAATCTTTAGAACAAAGATGATATGATCTGATCT 2858
QY 2705 GATGTTATATATTAATAATCTCAAAATTAAGAAATATGCTTCTCAAAAA 2754
Db 2859 GATGTTATATATTAATAATCTCAAAATTAAGAAATATGCTTCTCAAAAA 2908

RESULT 3
US-10-051-902-1
; Sequence 1, Application US/10051902

Db 434 GGGCTGACTGCGTTGGTAGAGAGCCCATGCTGCTCGCCTCGGCTGTCTCTACTTGTCA 493
QY 433 GCGGCTCATCATGCTATGCTCTCTAATGCTATGCTGCTGCTGGCAGCCTTGTAG 492
Db 494 GTGGGCTGTGATGCTTTGGCCGCAATTGTGTACATCTTCTCCTCGCAAGGCTCATTG 553
QY 493 ATGGAATTTGGTATTGGCTTGGCTGTACGCTTGTGCTTTGTACATTTCAAGAAATAGCCC 552
Db 554 ATGGGTTGGTATCGGTTGGCCGGTCAACACTGTCTCTCTACATCTCCGAACCTGCAC 613
QY 553 C---TTGGAGATTAGAGTTTGTCTGAATACACTACCAATTCAGTGG--ATCAGGAG 606
Db 614 CGCACAGANATTCTTGGGCTGNTNGAACACGTTGCCGAGTTCATTGGGGTCAGNGAG 673
QY 607 GAATGTTCTTGTACATCTGCAATGCTGTTGGAGATGCCCTGTGCCATCACCCGATTGGA 666
Db 674 GGATGTTCTCTCTACTGCAATGGTGTGGGATGTCCCTCATGCCCAACCTGATTGGA 733
QY 667 GAATTATGCTTGTGCTGCTCGCATACCTTCATTGTTCTTCTTGGTTGACAAATTTT 726
Db 734 GGCTCATGCTTGGAGTTCTGTGATCCCGTCACTTATNTACTTTGGACTGACTGTCTTCT 793
QY 727 ATCTTCCTGAATCTCCAAGATGCTCTGTAGCAAAAGTCCGATGGCAGAGGCAAAAAGG 786
Db 794 ACTTGCTGAATCACCAAGGTGCTGTGAGCAAAAGAGATGGCGAGGCGAAGAGAG 853
QY 787 TGTGCAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATTGTCCCTTCTTCTGCAAG 846
Db 854 TGNTGCAAAAGCTGCGGGGAGAGAAAGATGTCTCANGGAGANGGCTCTTCTAGTTGAAG 913
QY 847 GGTGAGGTTGAGAGACACTTCCATTGAAGATACATTTGACCTGCGACCGAGG 906
Db 914 GTTGGGGGTCGGTAAGAATACAGTATTNAGATACATTTGACCTGCGACCGAGG 973
QY 907 CAGCCGATGATCTGTTTACACGGTGATTAAGGAACAATCACACTTTATGGGCTTGAAG 966
Db 974 CAGCCGATGATCTGTTTACACGGTGATTAAGGAACAATCACACTTTATGGGCTTGAAG 1033
QY 967 AAGGCCAGTCATGATTGCTGCACTTCTAAGGGAGCCCATCATGCTTGAAGTGTGCTTT 1026
Db 1034 AAGGCCAGTCATGATTGCTGCACTTCTAAGGGAGCCCATCATGCTTGAAGTGTGCTTT 1093
QY 1027 CTCTGCACTCTGTCATGAGGAGCATGTTGAACCAAGTGTACCCCTTATGATCCGATTG 1086
Db 1094 CTCTGCACTCTGTCATGAGGAGCATGTTGAACCAAGTGTACCCCTTATGATCCGATTG 1153
QY 1087 TGACACTTTTGTAGTGTCCATGAGATATATGCTCAAGCTGGAAGAGATATGAGAGCA 1146
Db 1154 TGACACTTTTGTAGTGTCCATGAGATATATGCTCAAGCTGGAAGAGATATGAGAGCA 1213
QY 1147 CATTTGTTCCAACTTTGGAAGTATGTTCAAGTGTCAAGATCAGATCCGCAAAATGAGC 1206
Db 1214 CATTTGTTCCAACTTTGGAAGTATGTTCAAGTGTCAAGATCAGATCCGCAAAATGAGC 1273
QY 1207 AGTGGATGAAGAGATCTTCATAGGGATGACGAGAGTACGATCTGATGTTGACGAG 1266
Db 1274 AGTGGATGAAGAGATCTTCATAGGGATGACGAGAGTACGATCTGATGTTGACGAG 1333
QY 1267 GTGACTATGAGACAAATCTCCATAGCCCATTTGCTGTCCAGGCAAGCAAGGTGCGGAG 1326
Db 1334 GTGACTATGAGACAAATCTCCATAGCCCATTTGCTGTCCAGGCAAGCAAGGTGCGGAG 1393
QY 1327 GGAAGGACATTTGTGCAACATGCTCACCGTGAAGTGTGAGCATGAGAAGGCAAAACC 1386
Db 1394 GGAAGGACATTTGTGCAACATGCTCACCGTGAAGTGTGAGCATGAGAAGGCAAAACC 1453
QY 1387 TCTTAGGGAGGCTGAGATGCTGTGAGCAGCAGCTGATATCGTGGGGGATGCGAGCTTG 1446
Db 1454 TCTTAGGGAGGCTGAGATGCTGTGAGCAGCAGCTGATATCGTGGGGGATGCGAGCTTG 1513
QY 1447 CTTGAAATGCTCAGAGAGAGAGGTGAGATGTTAGAAAGAGAGGTGTTCAAAAGAG 1506

Db 1514 CTTGAAATGCTCAGAGAGAGAGGTGAGATGCTAGAAAGAGAGGTGTTCAAAAGAG 1573
QY 1507 TCTACTTGACCAAGAGAGGAGTTCCTGGCTCAGAAAGGGGCTCAATTGTTTCACTTCCCG 1566
Db 1574 TCTACTTGACCAAGAGAGGAGTTCCTGGCTCAGAAAGGGGCTCAATTGTTTCACTTCCCG 1633
QY 1567 GTGGTGGGATGTTTGTAGGGTAGTAGATTGTACATGCTGCTGCTTTAGTAAGTCAGT 1626
Db 1634 GTGGTGGGATGTTTGTAGGGTAGTAGATTGTACATGCTGCTGCTTTAGTAAGTCAGT 1693
QY 1627 CAGCACTTTTCTCAAAAGGCTCTTGCTGAACCAAGCATGTCAAGATGCTGCCATGTTCAAC 1686
Db 1694 CAGCACTTTTCTCAAAAGGCTCTTGCTGAACCAAGCATGTCAAGATGCTGCCATGTTCAAC 1753
QY 1687 CATCTGAGTAGCTGCCAAAGGTTCAACGTTGGAAGATTGTTGAACTGGAAGTGAAGC 1746
Db 1754 CATCTGAGTAGCTGCCAAAGGTTCAACGTTGGAAGATTGTTGAACTGGAAGTGAAGC 1813
QY 1747 GTGCCCTGTAGTCCGTTGGAATTCAAGATCTTCAACAGTTTGTGGAATAACGGTG 1806
Db 1814 GTGCCCTGTAGTCCGTTGGAATTCAAGATCTTCAACAGTTTGTGGAATAACGGTG 1873
QY 1807 TTCTGTACTATACCCCAAAATTTCTGAGCAAGCTGGTGGCAGTTATTTTCCAAAT 1866
Db 1874 TTCTGTACTATACCCCAAAATTTCTGAGCAAGCTGGTGGCAGTTATTTTCCAAAT 1933
QY 1867 TTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCTCTCACTACTTAATATGC 1926
Db 1934 TTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCTCTCACTACTTAATATGC 1993
QY 1927 TTCCTTGCAATTGGCTTTGGCCATGCTGCTTATGATCTTCCGGAAGAGGTTTGTGCTGC 1986
Db 1994 TTCCTTGCAATTGGCTTTGGCCATGCTGCTTATGATCTTCCGGAAGAGGTTTGTGCTGC 2053
QY 1987 TAGGCACAATTCGAATCTGTATAGCATCTCTAGTTATCTCTGTTGTGCCAATCTAATTG 2046
Db 2054 TAGGCACAATTCGAATCTGTATAGCATCTCTAGTTATCTCTGTTGTGCCAATCTAATTG 2113
QY 2047 ATTTGGTACACTAGCCCATGCTTGTCTCCACCGTCAAGTGTATCGTCTACTTCTGCT 2106
Db 2114 ATTTGGTACACTAGCCCATGCTTGTCTCTCCACCATCAGTGTATCGTCTACTTCTGCT 2173
QY 2107 GCTTCGTTATGGGATTTGGTCCCATCCCAACAATTTATGTGACAGAGATCTTTCCAACCA 2166
Db 2174 GCTTCGTTATGGGATTTGGTCCCATCCCAACAATTTATGTGACAGAGATCTTTCCAACCA 2233
QY 2167 GGGTTCGTGGCTCTGTATGTCATTTGTGCTTGTACATTTCTGATCGAGATATCATCG 2226
Db 2234 GGGTTCGTGGCTCTGTATGTCATTTGTGCTTGTACATTTCTGATCGAGATATCATCG 2293
QY 2227 TCACCTACAGCCTTCTGTGATGCTGAATGCTATTGGAAGTGGCGGAGTTTTCAGCATAT 2286
Db 2294 TCACCTACAGCCTTCTGTGATGCTGAATGCTATTGGAAGTGGCGGAGTTTTCAGCATAT 2353
QY 2287 ATGCAGTGTATGCTTATTTCTTGTGTTGCTTCTTAAGTCCCTGAGACAAAGG 2346
Db 2354 ATGCAGTGTATGCTTATTTCTTGTGTTGCTTCTTAAGTCCCTGAGACAAAGG 2413
QY 2347 GGATGCCCCCTGAGGTTATTAACCGAATTTCTTGCAGTTGGTGCAGAGCAAGCGCTGCAA 2406
Db 2414 GGATGCCCCCTGAGGTTATTAACCGAATTTCTTGCAGTTGGTGCAGAGCAAGCGCTGCAA 2473
QY 2407 AAGCCTAATTTCTTGGTACCTTTGTGTGCAACTATTGCACTGTAAATTAGAACTTGAA 2466
Db 2474 AAGCCTAATTTCTTGGTACCTTTGTGTGCAACTATTGCACTGTAAATTAGAACTTGAA 2533
QY 2467 GGGGTTTACCAAGAGCTCGAGAAATTACTTTGATTTGTGTAAATGTAAAGGAAACGA 2526
Db 2534 GGGGTTTACCAAGAGCTCGAGAAATTACTTTGATTTGTGTAAATGTAAAGGAAACGA 2593
QY 2527 ACATCTGCTCATGCTCTCAAAAGGTAAAGAGTCCCTCAATGCAAAATAGAGTCTGT 2586
Db 2594 ACATCTGCTCATGCTCTCAAAAGGTAAAGAGTCCCTCAATGCAAAATAGAGTCTGT 2653

QY 2587 TAAGTTGTCATGTCATTACCATATGTTTACCTATTGTACTGTATTATTAAGTCAAGC 2646
Db 2654 TAAGTTGTCATGTCATTACCATATGTTTACCTATTGTACTGTATTATTAAGTCAAGC 2713
QY 2647 TATTCACGCTGCTGTTGCTAGAAATCTTTAGAACAAAGATGATTAATGATCTGATCTGA 2706
Db 2714 TATTCACGCTGCTGTTGCTAGAAATCTTTAGAACAAAGATGATTAATGATCTGATCTGA 2773
QY 2707 TGTATTAATATTCAAATCTCAATTAAGAAATATCGTTTCTCAAAAAA 2757
Db 2774 TGTATTAATATTCAAATCTCAATTAAGAAATATCGTTTCTCAAAAAA 2824

RESULT 5
US-10-051-902-7
; Sequence 7, Application US/10051902
; Publication No: US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-051-902-7

Query Match 33.9%; Score 940.2; DB 14; Length 2601;
Best Local Similarity 66.0%; Pred. No. 2e-250;
Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;
QY 177 ATGTCGGGGCTGTTCTGTCGCCATAGTCGCCCTCCATCGGCAATCTATTGAGGGTGG 236
Db 175 ATGAAAGTGCCGCTGTTGTCATGTCGCCCTCCATCGGCAATCTATTGAGGGTGG 234
QY 237 GAGATGCAACCATCGAGCTGCTCTGTATATTAAGAAGAAATTTCAATTGCAAAAT 296
Db 235 GATTAATGCTACCATCGCCGGGCTAATGTTACATTAAGAAGAACTTGCTTGGGAA-- 292
QY 297 GAGCCCACTGTGAGGAGCTAATGTCATGTCACCTTATGCGGCCCAACCATCGTACT 356
Db 293 ----CAACTATGAAAAGGCTTGTTGGGCAATGTCCTGATTGAGCAACGGTAATCAC 348
QY 357 ACATTTCCGGGCCATTATCAGACTGATTGGCCGACGCCCTATGCTTATTCTCTTCA 416
Db 349 ACATGCTCTGCTCTATAGCGGATGCTCGGTGCGGACCAATGATTAATCTCATCT 408
QY 417 ATTCTGACTTCTTCAAGCGGCTCATGCTATGCTCTCTAATGCTATGCTCTGCTG 476
Db 409 GTGCTCTATTCTTGGGTGTTGGTGTGATGCTGTGCCCAAAATGTGATGTGTGTC 468
QY 477 TTGGACGCTTCGTAGATGATTTGGATTTGGCTGTGCAAGCTTGTGCTTGTAC 536
Db 469 TTGGGAGGCTACTTGATGATTTGGATTTGGCTTGTGCTGTGCTTGTCCCGGTCTAT 528
QY 537 ATTTCAGAAATAGCCCCCTTCGAGATTAAGGTTTGTGTAATACATAACCAATTCAGT 596
Db 529 ATAATGAAACGCGCGCTGTAATAAAGGGGCTGTGTAATACGCTTCTCAGTTTCACT 588
QY 597 GGATCAGAGGAATGTTCTTGTCTAATGCTATGCTGTGTTGGATGTCCCTGTGCCATCA 656
Db 589 GGCTCTGAGGAATGTTTGTCTAATGCTATGCTTGTGCAATGCTATGATGCCCGC 648

QY 657 CCCGATTGAGATATATGCTTGTGTGCTCGCGATACCTTCATGTTCTTCTTGGTTTG 716
Db 649 CCTAGCTGAGGCTATGCTTGGGTTCTGTCTATCTCTCTCTGTATTTTGCATTG 708
QY 717 ACAATATTTATCTTCTGTAATCTCCAGATGCTGTTAGCAAAAGCTCGCATGGCAGAG 776
Db 709 ACCATTTTCTTCTGCGAGTCTCTCGGTGCTGTGACCAAAAGAGATGCTCGAG 768
QY 777 GCAAAAAAGTGTGCAAAAGTTACGGGGGAAAGCATGCTCAGGTGAATTGCTCCTT 836
Db 769 GCTAAGAAAGTGTCTCAAAAGATTGCGGAGGAGAGATGTCTCAGGCGCATGCTTG 828
QY 837 CTTCGAAAGGTTGAGGTTGAGAGACACTTCCATTGAAGAGTACATCATTTGACCT 896
Db 829 CTGTTGAAGTCTCGGATTTGGGGTGTATCATCTATCGAAGAGTACATTAATTGGCCCT 888
QY 897 GCCACCGAGGACCGCATGATCTTGTACTGACGGTATAGGAACAATTCACCTTTAT 956
Db 889 GCTGACGATGTGCTGATGTCATGACATGCAACAGAAAGATTAATTCGATTATAT 948
QY 957 GGGCTGAAGAGGCCAGTCATGATGCTGCACTTCTAAGGAGCCCATCATGCTTGA 1016
Db 949 GGATCCCAAGCAGGCTTCTTGTGTTATCAAAACCTGTCACTGACAGAGTTCTATTGG- 1007
QY 1017 AGTGCTTCTCTTGCACTCTGTCATGAGCATGTTGAACCAAGATGATACCCCTTATG 1076
Db 1008 -----CCTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
QY 1077 GATCCGATTTGACACTTTTGTAGTGTCCATGAGAATATGCTCA-----AGCTGA 1130
Db 1057 GATCCTGTGTGACACTGTTTGTAGCATTCATGAGAAGCTCCCCGAGACAGACAAGA 1116
QY 1131 GGAAGTATGAGAGCACTTGTTCCAAACTTGAAGTATGTTCACTGTCACAGATCAG 1190
Db 1117 GGAAGCATGCGAAGCACTCTGTTCCAAATTTGAAGCATGTTCAAGCATGCTGAGCCG 1176
QY 1191 CATGCCAAAATGACAGATGGGATGAAGAAATCTTCATAGGATGACGAGATACGCA 1250
Db 1177 CATGCTAAATTTGAACAATGGGATGAGAAGCTTACAAAGGGAACGTGAGACTACATG 1236
QY 1251 TCTGATGTTGACAGAGTGAATATGAGACAACTTCCATAGCCCATGCTGTCCAGGAG 1310
Db 1237 TCAGATGCAACCCGTTGGGAGACTCCGATGATTAATTTGACACAGTCTTTAATCTCAAGCCAA 1296
QY 1311 GCAACAGTGCAGGAGGAAGACATTTGACCATGTCACCCGTGAAGTCTTTGAGC 1370
Db 1297 ACAACAAGCTTGAAGAAAGACTTACCTCTCTCTTCCATGCGAGATCTTGGCAGC 1356
QY 1371 ATGAGAAGCA---AACCTCTTAGGGGAGGAGTGTGAGCAGCACTGATATC 1427
Db 1357 ATGAGGCTCACAGTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGTAGTACAGTATT 1416
QY 1428 GGTGGGGATGGACGCTTGTGGAATGCTCAGAGAAGAGGTGAGAAATGCTAGAAAG 1487
Db 1417 GGTGTGCTGCACTGCACTGCAATGGAATGACGTGATAA---AGGTGAGATGGAACAA 1473
QY 1488 GAAAGTGTTCAAAGAGTCTACTTGCACCAAGAGGAGTCTGCTCAGAGAGGGGC 1547
Db 1474 CAAGGAGGTTTAAAGATTTATTACATGAGAGGAGGAGTTCTGCAATCTGCTGTGA 1533
QY 1548 TCAATGTTTCACTTCCCGGTGTGCGGATGTTTGAAGGTTAGTGAATTTGATCATGCT 1607
Db 1534 TCCATTTGATGATTCCTCGGTGAAGGCGAA-----TTTGTCCAGGCT 1575
QY 1608 GCTGCTTTAGTAAAGTCAAGCACTTTTCTCAAGAGGCTCTGCTGAACCAAGCATGTCA 1667
Db 1576 GCTGCTTTGTAAGCAACCCGCTCTTACTCCAGAGCTTATTGATGACACCAAGTT 1635
QY 1668 GATGCTGCATGTTCAACCATCTGAGGTAGCTGCCAAAGGTTCAAGTTGAAGATTTG 1727
Db 1636 GGGCTGCAATGTTCAACCATCTGAGACAGCTTCAAGGGCCAAAGTTGAAAGCTCTT 1695
QY 1728 TTTGAACCTGAGTGAAGCGCTGCCCTGTTAGTGGTGTGGAATTCAAGATCTTCAACAG 1787

Db	1696	CTTGAACACGAGGGTTAAGCATGTCATGGTGTGAGTTGGAATACAAATACTTCAGAG	1755
QY	1788	TTTGCTGAATAAACGGTGTCTGACTATATACCCACAATCTTGAGCAAGCTGTGTG	1847
Db	1756	TTTTCAGGAGTAATAGGGTCTATATTAACACACCTCAATCTTGAAGAGGCCGGTGT	1815
QY	1848	GCAATTATCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCT	1907
Db	1816	GAAATTCTTCTTCAATATAGGCAATTGGCTCAGAGTCGGCATCAATCTTATCAGTCT	1875
QY	1908	CTCAGTACCTTACTAATGCTTCTTGCAATGGCTTGGCAATGCTGCTTATGATCTTCC	1967
Db	1876	TTCAACAACCTTCTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGATGTTCA	1935
QY	1968	GGAAGAAGTTTTTGTGCTGAGCACAATTCGAATCTTGAATAGCATCTAGTTAATCTG	2027
Db	1936	GCGAGAAGGACGTTGCTACTTACATACATCCCGTGTGATGTCATCTATATTTTG	1995
QY	2028	GTTGTGTCATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCACCGTCACT	2087
Db	1996	GTCATTTGAAGCCTGTGAATTTTGGCAATGTCGCCCATGACAAATCTCAACAGTATGC	2055
QY	2088	GTTATGCTCTACTCTGCTGCTTGTATGGGATTTGGTCCCATCCCAACATTTATGT	2147
Db	2056	GTTGTGTTTATTTCTGCTGCTTGTGATGGGTTATGACCAATTCGAACATCTTGTGC	2115
QY	2148	GCAGAGATCTTTCGAACAGGGTGTGCTGCTGTATGTCATTTGTGCTTACATTC	2207
Db	2116	TCAGAGATTTTCCCACTAGGGTGTGCTGCTGTATGTCATTTGTGATGTTCT	2175
QY	2208	TGATCGAGATATCATGCTCACTACAGCCTTCTGTGATGTCATTTGTGCTTACATTC	2267
Db	2176	TGATTTGAGACATCATCATCACTACTGCTGCTGTGATGCTGCTTATGACTT	2235
QY	2268	GCGGCTGTTTTCAGCATATATGAGTGTATGCTTGTATTTCTTGTGCTTCTT	2327
Db	2236	GCTGTGTATTCGCCATTTACGAGTGTGTTTGTTCATCTGCGATATTTGTGTTTG	2295
QY	2328	AAGTCCCTGAGACAAAGGGATGCCCTTGAGGTTATACCAATTTCTTGCAGTTGGT	2387
Db	2296	AAGTTTCCAGAAACAAAGGGATGCCCTTGAAGTCACTCTGAATTTCTTGTGTGA	2355
QY	2388	GCGAAGCAAGCGCTGC	2404
Db	2356	GCAAGCAGGCTGCTT	2372

RESULT 6
US-10-051-909-7

; Sequence 7, Application US/10051909
; Publication No. US20020199217A1

; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim

; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B1163 US CIP

; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17

; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7

; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max

US-10-051-909-7
Query Match 33.9%; Score 940.2; DB 14; Length 2601;

QY	177	ATGTCGGGGCTGTCTTGTCCCATAGTCCCTCCATCCGCAATCTATGAGGGGTG	236
Db	175	ATGAAGGTGCCGCTCTTGTGTCTATTTGCCCTTCCATTTGTAATTTCTCCAGGATGG	234
QY	237	GACATGCCACCATCGAGCTGCTGTCTGTATTAAGAAGAAATTTCAATTGCCAAAT	296
Db	235	GATATGCTACCATCGCGGGGCTAATGTTACATTAAGAAGAACCTTGCTTGGGAA--	292
QY	297	GAGCCACTGTGAGGCACTAATTGTGTCATGCTTATCGCGCCACCATCTTACT	356
Db	293	---CAACTATGAAGGCTTGTGTGAGCATGTCCTGATTGAGCAACGGTAATCAC	348
QY	357	ACATTTCCGGCCATATACAGCTCGATTGGCCGACGCCCTATGCTTATCTCTTCA	416
Db	349	ACATGCTGTGCTTATAGCGGATTTGGCTCGGTCCGACCATGATGATATCTCATCT	408
QY	417	ATTCTGACTTCTTACCGGCTCATCATGCTGCTCTCTAATGCTATGCTGCTGCTG	476
Db	409	GTCCTTATTTCTTGGGTGTTTGTGATGCTGTGTGCTCCCAATGTGTATGTGTGCTG	468
QY	477	TTGGCAGCTTCTGATGATTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	536
Db	469	TTGGCAGGCTACTGATGATTTGATTTGGATTTGCTGCTGCTGCTGCTGCTGCTGCT	528
QY	537	ATTTCAGAAATAGCCCTTCGAGATTAGAGTTGCTGTAATACACTACCAATTCAGT	596
Db	529	ATATCTGAACCGCGCTCTGAATTAAGGGGCTGTGAATACGCTTCTCAGTTCACT	588
QY	597	GGATCAGAGGAATTTCTGTCTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	656
Db	589	GGCTCTGAGGAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	648
QY	657	CCCGATTGAGAAATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	716
Db	649	CCTAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	708
QY	717	ACAATATTTATCTTCTGTAATCTCCAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	776
Db	709	ACCATTTTCTTCTGCGAGTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	768
QY	777	GCAAAAAGGTGTGCAAAAGTTACGGGGGAAAGCATGTCTCAGTGAATGTCCCTT	836
Db	769	GCTAAGAAGGTGCTCAAGAATTCGCGGAAGGAGATGTGTGAGGAGATGCTGCTGCTG	828
QY	837	CTTCTGAGAGGTTGAGGTTGAGAGACATTCATTTGAAGATACATCATTTGACCT	896
Db	829	CTGTTGAAGTCTCGGATTTGGGGTATATCATCTATCGAAGATACATATTTGGCCT	888
QY	897	GCCACGAGGACCGCATGCTTGTACTGACGGTGATTAAGAACAAATCACATTTAT	956
Db	889	GCTGACGATGTGCTGATGCTCATGAACATGCAACAGAAAGATAAATTCATTTATAT	948
QY	957	GCGCTGAAGAAGGCAATCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1016
Db	949	GGATCCCAAGCAGGCTTCTTGTGTTATCAAACTGTCACTGACAGAGTTCTATTTG-	1007
QY	1017	AGTGTCTTCTTGTGATCTGCTCATGAGGAGCATGTGTAACAGAGTGTACCCCTATG	1076
Db	1008	-----CCTTGGCTCACACCATGGAAGCATCAACCAAGCATGCCCTCATG	1056
QY	1077	GATCCGATTTGACATTTTGTGATGCTCATGAGATATGCTCA-----AGCTGA	1130
Db	1057	GATCCTCTGTGACACTTTTGTGATGATCATGAGAGCTCCCGAGACAGAGCAAGA	1116
QY	1131	GGAATATGAGAGCATTTGTTCCAACTTTGGAAGTATGTTCACTGTCAGATCAG	1190
Db	1117	GGAAGCATGCGAAGCACTCTGTTCCAAATTTGGAAGCATGTTCAAGCATGCTGAGCCG	1176
QY	1191	CATGCCAAAATGAGCATGGGATGAAGAAATCTTATAGGATGACGAGAGTACGCA	1250

Db 1177 CATGCTAAATTTGAACAATGGGATGAAGAAGCTTACAAGGGAACGTGAGACTACATG 1236
QY 1251 TCTGATGTCAGAGAGTGAATATGAGACAATCTCCATAGCCCATTTGCTGTCCAGCAG 1310
Db 1237 TCAGATGCAACCCCGTGGGACTCCGATGATTAATTTGCACAGTCTTAACTCTCACGCCAA 1296
QY 1311 GCAACAGGTGCGGAAGGGAAGGACATTTGTGCACCATGTGTACCCGTGGAAGTCTTTGAGC 1370
Db 1297 ACAACAAGCCTTGAAAAAGACTTAACCTCTCTCTCCATGGCAGTATCTTGCGAGC 1356
QY 1371 ATGAGAAGCA--AACCTCTTAGGGGAGGGTGAATGTGTGAGCAGCAGCTGATATC 1427
Db 1357 ATGAGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGTAGTACAGGTATT 1416
QY 1428 GGTGGGAGTGGCAGCTTGTGAAATGTGTGAGAGAAGAGTGAATGTGTAGAAG 1487
Db 1417 GGTGGTGGCTGCAACTGGCATGGAATGACTGATAA---AGGTGAGATGGAACAACAA 1473
QY 1488 GAAGGTGTTTCAAAAAGTCTACTTGACACCAAGAGGAGTTCTGCTCAAGAGGGC 1547
Db 1474 CAAGGAGGTTTAAAGATTTATTTACATGAGAGAGGAGTTTCTGCATCTCGTGTGA 1533
QY 1548 TCAATTTCTTCACTTCCCGGTGTGGCGATGTTTGAAGGTAGTGAAGTTTGTACATGCT 1607
Db 1534 TCCATTGTATCGATTCCCGGTGAAGCGCAA-----TTGTGTCAAGGCT 1575
QY 1608 GCTGCTTTAGTAGTCACTGACGACATTTTCTCAAAAGGCTGTGCTGAACCAAGCAGTGTCA 1667
Db 1576 GCTGCTTGTGTAAGCCACCCGCTCTTACTCCAAGAGCTTATGTGAGACACCAAGTT 1635
QY 1668 GATGCTGCCATGTTCAACCACTGTAGGTAGCTGCAAGGTTCAAGTTGAAGATTGTG 1727
Db 1636 GGGCTTGAATGTTTCAACCACTGTAGGTAGCTGCAAGGTTCAAGTTGAAGATTGTG 1695
QY 1728 TTTGAACCTGAGTGAAGGCGCTGCTGTAGTGTGGAATTGAGATCTTGAACAG 1787
Db 1696 CTTGAACCAAGGTTAAGCATGATGTTGTTGAGTGTGAATACAAATACTTCAAGCAG 1755
QY 1788 TTTGCTGGAATAACCGTGTCTGTACTATATACCCCAAAATCTTGAGCAAGCTGTG 1847
Db 1756 TTTTCAGGATTAATGGGTTCTATATTACACACCTCAAACTTGAAGAGGCGCGTGT 1815
QY 1848 GCAATTAATCTTCCAAATTTGCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCT 1907
Db 1816 GAAGTTCTTCTTCAATATAGGCAATGCTCAGAGTGGCATCATCTTATCATGCTCT 1875
QY 1908 CTCACTACCTTACTAATGCTTCTGTCATGCTGCTTGGCATGCTGCTTATGATCTTCC 1967
Db 1876 TTCAACAACCTTCTGATGCTTCCCTGTATAGGCGTAGGCATGAAGCTCATGATGTTCA 1935
QY 1968 GGAAGAAGGTTTGTGCTGAGGCACAATTCCAATCTGATAGCATCTCTAGTTATCTCTG 2027
Db 1936 GGCAGAAGGAGTGTGCTACTTAACAATCCCCGTGCTGATGTGTCATCATTAATTTG 1995
QY 2028 GTTGTGTCATTAATTAATTTGGGTACACTAGCCCATGCTTGTCTCCACCGTCACT 2087
Db 1996 GTCAATTGGAAGCTGTGAATTTTGGCAATGTGCGCCCATGCAACAATCTCAACAGTATGC 2055
QY 2088 GTTATCGTCTACTTCTGCTGCTTGTATGAGATTGTGTCATCCCAACATTTATGT 2147
Db 2056 GTTGTGTTTATTTCTGCTGCTTGTGATGGGTATGACCAATTCMAACATCTTTGTC 2115
QY 2148 GCAGAGATCTTTCCAACGAGGTTGCTGCGCTGTATGTCATTTGTGCTTACATTC 2207
Db 2116 TCAGAGATTTTCCCACTAGGCTGCTGCGCTGTGATGCTATCTGTGATTAATGTTTC 2175
QY 2208 TGGATCGAGATATCATGTCACCTACAGCCTTCTGATGCTGAATGCTATGAGCTG 2267
Db 2176 TGGATGAGACATCATCATCACTACTGCTGCTGATGCTCGGCTCTTTAGAGCTT 2235
QY 2268 GCGGCTGTTTCAAGCATATGACAGTGTGATGCTTGAATTTCTTGTGCTGCTTCTT 2327
Db 2236 GGTGCTGTATTCGCCATTTACGAGTGTGTTGTTTCACTCTGAGATATTTGTGTTTG 2295

QY 2328 AAGTCCCTGAGACAAGGGGATGCCCCCTTGAGGTATTACCGAATCTTTGCAAGTGTG 2387
Db 2296 AAGTTCCAGAAACAAGGGCATGCCCCCTTGAAGTCACTCTGAATTTCTTGTGGA 2355
QY 2388 GCGAAGCAAGCGGCTGC 2404
Db 2356 GCAAGCAGGCTGCTC 2372

RESULT 7
US-09-938-842A-1315
; Sequence 1315, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1315
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1315

Query Match 31.5%; Score 875.8; DB 10; Length 2190;
Best Local Similarity 64.9%; Pred. No. 1.5e-232;
Matches 1449; Conservative 0; Mismatches 722; Indels 60; Gaps 8;

QY 177 ATGTGGGGCTGTCTTGTGCGCATAGTGCCTCCATCGCAATCTATTGACAGGGTGG 236
Db 1 ATGAGTGAAGCTGTGCTTGTGCTATGCTGTGCTGTGCAACTGTTTACAAGATGG 60
QY 237 GACAATGCCACCATCGCAGCTGCTGTTCTGTATATAAAGAAGAAATTTCAATGCAAAAT 296
Db 61 GATAACGCAACTATTGCAAGAGAGCTGTGTGATCAATAAAGAGATTATTTGAGAGT 120
QY 297 GAGCCCACTGTGAGGGAATAATTGTGTCATGTCATTATCGGCGCCACCATGCTTACT 356
Db 121 AATCATCAGTGAAGGTTCTAATTGTGCGCATGTCACTTATGTGTACTCTGATTACA 180
QY 357 ACATTTCCGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTTCA 416
Db 181 ACATGCTTGAAGGAGTACGTGATTGGCTGTGCGCCGTCCCATGCTTAATTTGCTCA 240
QY 417 ATTCTGACTTCTTCAAGCGGCTCATCATGCTATGCTCTCTAATGCTATGCTGCTG 476
Db 241 ATTCTGACTTGTGTGTTCTCTAGTAATGCTATGCTCTCGAATGTTATGTGTGCTC 300
QY 477 TTGCACGCTTGTGATGATTTGTTGCTTGGCTGTCAAGCTGTGCTTGTGAC 536
Db 301 TTAGAAGGTTGTAGATGATTTGGGGTGTCTGTGTGTCACACTGTCTTCTATTAT 360
QY 537 ATTCAAGAAATAGCCCCCTTGGAGATTAGAGTTTGTGTAATACATACATCAATTCAGT 596
Db 361 ATATCTGAGACTGCAACCACTGAGATTAGGGGAGCTGTGAATACGCTTACCGCAGTTC 420
QY 597 GATCAGAGAGAAATGTTCTGTCACTACTGATGCTGTTGGAGATGCTCTGTGCGCATCA 656
Db 421 GGTCTGAGAGGATGTTCTTATCTTACTGATGATGTTTTCGAATGCTGTGATGCTATCA 480

QY	657	CCCGATTGGA	AATTA	TGCTTG	TGCTG	CTCGC	GAATAC	CTTCA	TTGTTCTT	CGTTG	716
Db	481	CCTAGCTGAG	ATTGA	TGCTTG	TGCTT	CTTTC	ATTCCTT	CCCTT	GTCTTT	CTTCTC	540
QY	717	ACATATTTT	ATCTTC	GAATCT	CCAAGAT	GGCTGT	AGCAAA	GGTCG	GAATG	GCAGAG	776
Db	541	ACGCTCTT	CTCTTG	CCCGAG	TCCCCA	AGGTGG	CTGTGA	GCAAA	AGSTCG	AATGCTTGA	600
QY	777	GCAAAAAG	STGTG	CAAAAG	TTA	CGGGGA	AAACG	ATGCT	CAG	GTGAATG	836
Db	601	GCAAA	CGGCTT	CTTC	AGAC	ACTG	CGTG	TCGCG	AGAC	ATGTGCTG	660
QY	837	CTTCTG	AAGG	TTG	AGGTTG	AGAGAC	ACTTCC	ATTGA	AGATCA	TGGA	896
Db	661	TTGTTG	AGGGT	CTTG	GAATG	GAGTGA	AAAC	ATAG	AGAA	TATATATG	720
QY	897	GCCACC	GAGC	AGCCG	ATGATCT	TGTTACT	GACGGT	GTAA	GAACA	ATCACA	956
Db	721	GCGATGA	AGTTAC	TGATG	ATCATG	ATATAG	CTGTG	GAATA	AGATCA	AATTA	780
QY	957	GGGCTGA	AGAGG	CCAGTCA	TGATG	TTGCT	CGACCTT	CTA	AGGAC	CCATCA	1016
Db	781	GGTGCA	GAAGA	GGGCTG	AGTTGG	GTGCT	AGGCC	AGTCA	AAAG	-----GGA	828
QY	1017	AGTGTCT	TTCTCT	TGATCT	CTGTC	ATGGAG	CAATG	GTGA	CCAG	-----TGTA	1073
Db	829	AGCACT	ATGAGT	GTTTGT	CTCGC	ATGGA	ATCAATG	AGCAG	AGGCA	AGGCTCAT	888
QY	1074	ATGATCC	GATTG	TGAC	CTTTT	TGTTG	TAGTGT	CCATG	AAATAT	GCTCA	1133
Db	889	ATTGAT	CTCTT	GTCA	CACTG	TTGGAG	CGTTCA	GAGA	AGATG	CCGAC	945
QY	1134	AGTATG	AGGAC	ATTTG	TTCCAA	CTTTG	GAATAT	GTTC	AGTGT	CA	1193
Db	946	AGCATG	AGAGT	GCCTT	GTTC	CCACAT	TTTGG	AGATG	TTCA	GTGAGG	1005
QY	1194	GCCAAA	ATGAG	CAGT	GGATGA	AGAA	ATCTT	CATAG	GGATG	ACGAG	1253
Db	1006	CCAAGA	CATGA	ATTGG	ATGA	AGAA	ATCTT	GTG	AGAA	GGTGA	1065
QY	1254	GATGTG	CAGAG	TGACT	ATAG	AGACA	ATCTC	ATAG	CCATTG	CTGCC	1313
Db	1066	GACCAT	---G	GAGAT	GTCTG	AAATG	ATCTCAT	TTCCG	TGATCT	CA	1122
QY	1314	ACAGTG	CGGA	AGGA	ACAT	TGTG	CA	CCATG	TCA	CCGTGA	1373
Db	1123	ACAAG	CATGA	---GA	AGAC	ATGCT	CTCAC	ATGGA	ACTCTT	CT	1179
QY	1374	AGAA	GC	AAAC	CTTTAG	GGGAG	GGGTG	AGATG	GTG	AGCAG	1433
Db	1180	CATG	AA	GTCA	AGTGA	GGAG	CTCA	AGGG	AGAC	GGGTAG	1239
QY	1434	GGATG	CAGC	TTGCTT	GGAATG	GTCA	GAAGA	GAAGTGA	ATG	GTAGAA	1493
Db	1240	GGATG	CAAGT	GCATG	GAATG	ACGGA	AAAGAA	GATGA	ATCG	GACAGAA	1299
QY	1494	GGTTT	CAAA	AGATCT	TGCA	CCAGA	GGAGTTC	TG	GGCTCA	GAAT	1553
Db	1300	GGTTT	-----	-----	-----	-----	-----	-----	-----	-----	1329
QY	1554	GTTTCA	CTTCC	GGTGG	TGCGAT	GTTTT	TGAGG	GTAG	AGTTG	TACATG	1613
Db	1330	GTTTCA	TTGCT	GTGTG	ATGGA	CCGGT	GAGCA	---G	ATTTG	TACA	1386
QY	1614	TTAGT	AGTCA	GTCA	CTTTCT	CAAA	GGGTCT	TGTA	ACCA	CGCATG	1673
Db	1387	TTGTTA	AGCA	ACAGCT	TTAT	TCCA	AAAC	CTTCA	AAAG	ATACAT	1446
QY	1674	GCCATG	GTTC	ACCCAT	CTGAG	GTAG	CTGCC	AAAGT	TCACG	TTGGA	1733
Db	1447	GCTATG	TACAT	CCATC	CGAA	CAACT	---AA	AGGTCA	ATTG	GATG	1503
QY	1734	CCTGAG	TGAGG	CGCTG	TGTTAG	TGCGT	GTGGA	TTCA	GATC	CTTCA	1793

Db	1504	CCTGAGTCAAGCGTGCATTAGTCGTAAGAGTTGGACTTCAATATCTTACAGCACTTCTCA	1563
QY	1794	GGATTAACGGGTGTTCTGTACTATACCCCAAAATTTCTGAGCAAGCTGGTGGCAGTT	1853
Db	1564	GGCATCAACGGAGTTCTTACTACACACCGCAAAATCCTTGAGCAGCGGGTGTGGGATC	1623
QY	1854	ATTCTTCCAAATTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACT	1913
Db	1624	CTACTATCGAACATGGGGATTAGTTCCTCCAGCATCCTTACTTATAGTGCAATTGACA	1683
QY	1914	ACCTTACTAATGCTTCCCTTGCAATGGGCTTGGCCATGCTGCTTATGGATCTTCCGAAGA	1973
Db	1684	ACCTTGTGATGTTAACCTGCAATAGCTGTGTGCAATGAGGCTCATGGATCTTCTGTCGA	1743
QY	1974	AGGTTTTCGTGCTAGGCAACAATTCATCTTGATAGCATCTCTAGTTATCCTGTTGTG	2033
Db	1744	AGGACCTGCTTCTCACCACGATACCAATCCTGATAGCATCTCTATGGTTTGTAGTAATC	1803
QY	2034	TCCAATCTAATTGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACGTCAGTGTATC	2093
Db	1804	TCAATCTGTGTTACATGACAGACGATTTGTGCAAGCGGTCTTATCAACCGTAAGCCTGTG	1863
QY	2094	GTTACTTTCGTGCTTCTGTTATGGGATTTGTTCCCATCCCAACAATTTATGTGCAGAG	2153
Db	1864	CTTACTTTCGTCTTCTGATGAGGTTTCGGTCTGCTCCAACATCTCTGTTCAGAG	1923
QY	2154	ATCTTTCCAACGAGGTTGCTGCGCTCTGTATGTGCCATTTGTGCTTACATTTCTGATC	2213
Db	1924	ATTTTTCCAACTCGAGTCCGCGAATCTGCATTCGCCATCTGCGCACTCACTTCTGATC	1983
QY	2214	GGAGATATCATTCGTACACCTACAGCCTTCTGTGATGCTGAATGCTATTGCACTGGCGGT	2273
Db	1984	TGTGACATATTCGTCACTTACAGTCTCCCGGTGCTGCTCAAAATCCATGGACTAGCTGT	2043
QY	2274	GTTTTCAGCATATATGCAGTGTGATGCTGATTTCTTGTGTTGCTTCTTCTTAAGGTC	2333
Db	2044	GTTTTCGAATGTACGCAATCGTATGTGCAATTTCAATGGGCTTGTGTTCATTTAAAGTC	2103
QY	2334	CCTGAGACAAAGGGATGCCCTTGAGGTTATTAACGAATTTCTTGAGTGGTGCAGAG	2393
Db	2104	CCGGAACCTAAAGGCATGCCACTTGAAGTCAACAAGAGTCTTTCTGTGGAGCTAGA	2163
QY	2394	CAAGCGGCTGC	2404
Db	2164	CAAGCTGAAGC	2174

```

RESULT 8
US-09-938-842A-2254
; Sequence 2254, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2254
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

```

US-09-938-842A-2254

Query Match 25.2%; Score 700; DB 10; Length 2205;
Best Local Similarity 61.2%; Pred. No. 1.4e-183;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;

QY 177 ATGTCGGGGGCTGTTCTTGTCCGCAATAGTCGCCCTCCATCGGCAATCTATTGCAGGGGTGG 236
DB 1 ATGAAAGGAGCGACTCTCGTTGCTCTCGCCGACACAATCGGCAATTTCTTAGAAGATGG 60
QY 237 GACAAATGCCACCATCGCAGCTGCTGTCTGTATATAAGAAAGAAATTTCAATTGCAGAAAT 296
DB 61 GACAAATGCCACCATTGTGAGCTATGGTTTATATCAACAAGAACTTGAATCTAACCA-- 118
QY 297 GAGCCCACTGTGAGGAGACTAATGTGTCAATGTCACTTAATCGGCCACCATCGTTACT 356
DB 119 ---CCTCTGTTCAAGGCTTTGCTGCTGTCTATGTCAATTGATCGGTGCAACGTCATCACG 174
QY 357 ACATTTCTCCGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTATTCTCTCTCA 416
DB 175 ACTTGCTCAGGACCGATATCTGATTTGGCTCGGCAGACGCCCATGCTCATTTATCATCA 234
QY 417 ATTCTGTACTTCTTCAGCGGCTCATCATGCTATGGTCTCTAATGTCTATGTCTGCTG 476
DB 235 GTTATGTATTTCGTCTGCGGTTTGATTAATGTGTGGTCTCCCAATGCTCATGTCTGTGC 294
QY 477 TTGGCAGCCTTCGTAGATGATTTGGTATTGGCTTGCTGTCAAGCTGTGCTTTGTAC 536
DB 295 TTGCTAGGCTTCTTAATGGGTTTGGTCCGGCTCGCGTTACACTTGTCCTGTTTAC 354
QY 537 ATTTCAGAAATAGCCCCCTTCGAGATTAGAGTTTGCTGAATACACTACCACAATTCAGT 596
DB 355 ATTCTGAACCGCTCTCCGAGATCAGAGACAGTTAAATACTCCCTCAGTTTCTT 414
QY 597 GGATCAGAGAGAAATGTTCTTGTCTACTACTGATGATGTTGGATGTCCTGTGCCATCA 656
DB 415 GGCTCTGTGGAATGTTTGTCTACTATGATGTTTCACTATGTCCTGATGACTCC 474
QY 657 CCCGATTGGAGAATTATGCTTGGTGTGCTCGGATACCTTCATGTCTTCTTGGTTG 716
DB 475 CCTAGCTGAGAGCCATGCTCGGTCTCTCGATCCCTTCTCTTATTGTTCTC 534
QY 717 ACAATATTTTATCTTCCCTGATCTCCAAGATGCTCTTAACAAAGTCGATGCGACAG 776
DB 535 ACGGTGTTTATTTGCCGAGTCTCTCGTTGGCTGTTAAGTAAAGGAAGATGACGAG 594
QY 777 GCAAAAAAGGTGTTGCAAAAGTTACGGGGAAAGACATGTCTCAGGTGAATGTCCCTT 836
DB 595 GCTAAGCAGTCTTCAACAGTTATGTGGCAGAGAAAGTATTACCGATGAGATGCTTTA 654
QY 837 CTTCTCGAAGGTTGAGGTTGAGAGACACTTCCATGAAGATACATCATTTGAACCT 896
DB 655 CTAGTTGAAGACTAGATATAGAGAGAGAAAAACAATGGAAGATCTCTTAGTAATTG 714
QY 897 GCCACCGAGCAGCCGATGAT--CTTGTACTGACGATGATAAGAAACAATCACACTT 953
DB 715 GAGATCATGAAGGTGATGATACACTTGAACCGTTGATGAGATGACAATGCGGCTT 774
QY 954 TATGGGCTGAAGAAGGCCATCATGATGCTCGACTTCTAAGGACCCCATCATGCTT 1013
DB 775 TATGGAACCCACGAGATCAATCGTACCTGTGCTAGACCTGTCCAGAACAAA----- 826
QY 1014 GGAAGTGTCTTCTCTTCATCTCGTCATGGGAGATGTTGAACCAAGTGTACCCCTT 1073
DB 827 -ATACTCACTTGGGCTAGGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATCCTT 885
QY 1074 ATGATCCGATTGTGACATTTTGTGTAGTGTCCATGAGAAATATGCTCAAGCTGAGGA 1133
DB 886 AAAGATCCGCTGTCAATCTTTTGGCAGTCTCCAGAGAAAGATGCCAAGCAGGCGGA 945
QY 1134 AGTATGAGAGACATGTTTCCAAACTTTGGAAGTATGTTAG-----TGTCAAGAT 1187
DB 946 AACACTCGAGTGGATTTTCCCTCATTTCCGAAGCATGTTCACTACTGCCGATGCG 1005

QY 1188 CAGCATGCCAAAAATGAGCAGTGGGATGAAGAG-----ATCTTCATAGGGATGAC 1238
DB 1006 CCTCAGGTAAACCGGCTCATTTGGAAAAAGACATAGAGAGCCATTACAAACAACAAT 1065
QY 1239 GAGAGTACGATCTGATGTGTGACAGAGGTGACTATGAG-----ACAATCTCCAT 1289
DB 1066 GATGACTATGCGACTGATGATGTGCGGGTGAATGATGACTCGGACAAGATTTGGCT 1125
QY 1290 AGCCCATGTCTGTCCAGCAGCAGCAAGGTGCGGAAGGAAGACATGTGCAACCATGCT 1349
DB 1126 AGCCCTTAATGTCCGGCCAGACCAACAGCATGGA---CAAGATATGATCCACATCCT 1182
QY 1350 CACCGTGAAGTCTTTGAGCATGAGAAAGCAAAACCCTCTTAGGGGAGGTGAGATGCT 1409
DB 1183 ACAAGTGAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCA---AGCCAACGGC 1239
QY 1410 GTGAGCAGCACTGATATCGGTGGGGATGAGCCTTGCTTGAATGTTCAGAGAAAGAA 1469
DB 1240 GAAAGTAGCATGGGAATTGTGTGTGGTGGCATATGGAT----- 1279
QY 1470 GGTGAGAAATGTAGAAAAGAGGTGTTTCAAAAAGTCTACTTGCACCAAGAGGAGTT 1529
DB 1280 -----ATAGATACGAAAAAGATGAATAAAGAGTATTATCTTAAAGAAAGATGA 1329
QY 1530 CCTGCTCAAGAGGGCTCAATTTGTTCACTTCCCGGTGTGGCGAGTTTTGAAGGT 1589
DB 1330 GCTGAATCTCGCGGTGCTCGATCATCTTAATCCCGAGAGTCCGATGCT-----GGA 1383
QY 1590 AGTAGTTGTACATGCTGCTGCTTATGTAAGTCACTCAGCAGCTTTCTCAAGGCTCTT 1649
DB 1384 GGCAGCTACATTCACGCTTCTGCCCCCTGTAAAGACAGATCTGTTCTTGGTCTTAAATCAGTT 1443
QY 1650 GCTGAACCAAGCATGTCAAGTCTGCCATGCTTCAACCATCTGAGTAGCTGCCAAAGGT 1709
DB 1444 CATGA-----TCCGCATGTGTTCCCGGAGAAAAATGTGCTGCTCTGGA 1488
QY 1710 TCAGTTGAAAGATTTGTTGAACTGAGTGAAGCGGTGCCCTGTTAGTGGGTGTTGA 1769
DB 1489 CCACTGTGCTGCTCTTCTTGAACTGTGTAAAGCGTGTGTTGTGTGTGCGGC 1548
QY 1770 ATTCAAGTCTTCAACAGTTTGTCTGGAATAAAGGTGTTCTGTACTAATACCCACAATTT 1829
DB 1549 ATTCAATATCTCAGACAGATTTTCAAGTATCAATGAGATTCTCTACTACACTCCTCAGATT 1608
QY 1830 CTTGAGCAAGCTGTGTGCGAGTTATTTCTTCCAATTTGCTCAGCTCGGCATCAGCA 1889
DB 1609 CTCGAACGGGCTGCGGTAGATATTTCTTTTTCGAGCCCTCGACCTAAGTTCATCTCTCGC 1668
QY 1890 TCCATCTGATGATTTCTCTCACTACCTTAATATGCTTCTTGCAATTTGGCTTTGCCATG 1949
DB 1669 TCATCTCTCATCAGCGGTTTAAACAACATTACTCATGTCTCCAGCCATTTGCTTGCATG 1728
QY 1950 CTGCTTATGATCTTTCGGGAAGAGTTTGTCTGCTAGGACAAATTCATCTTGATA 2009
DB 1729 AGACTCATGATGATATCCGGAAGAGTCAATTACTTCTCTGACAAATCCAGTTCTCATTT 1788
QY 2010 GCATCTTATGATATCTGTGTGTGTCATTAATGATTGGGTGACTAGCCCATGCT 2069
DB 1789 GTCTCACTTGTGCTCTTGTCACTCAGCGAGCTCATCCACATCAGCAAAAGTGTGAACGCA 1848
QY 2070 TTGCTTCCAGCGTCAAGTTATCGTTACTTCTGCTGCTTGTGTTAAGGATTTGTGCC 2129
DB 1849 GCATCTTCCAGAGTTGTGTGTCTACTTCTGCTTCTGTAATGGGTTACGCTCCC 1908
QY 2130 ATCCCAACATTTATGTGACAGAGATTTTCCAACCAAGGTTCTGTGCTCTGTATGCC 2189
DB 1909 ATTCCAACATCTCTGTCTGAAATCTTCCCAACAAGAGTCCGTGTCTGTGATGCC 1968
QY 2190 ATTGTGCTTTTACATTTCTGATCGAGATATCATGTCACCTTACAGCTTCTGTGATG 2249
DB 1969 ATATGTCTAATGTCTTTTGGATTGAGACATTAATGTACAGTACTCATCTCCGTTCTC 2028

QY 2250 CTGAATGCTATTGGACGTGGGGGTGTTTCAGCATATATGAGTCTGTATGCTTGAATTTCC 2309
Db 2029 CTCAGCTCGATCGAGTATGTTGGTGTTCAGCATTTAGCTGCGGTTGCGTTATCTCA 2088
QY 2310 TTTGTGTCGTCCTTCCCTTAAGGTCCTGAGACAAAGGGGATGCCCTTGAGGTATTAACC 2369
Db 2089 TGGATCTTCGTTTACATGAAGTCCCGAGACTAAAGGCATGCTTTGGAAGTTATCACA 2148
QY 2370 GAATTCCTTTGCAAGTTGGTGC 2389
Db 2149 GACTACTTTGCTTGGAGC 2168

RESULT 9

US-10-051-902-13
; Sequence 13, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-13

Query Match 24.9%; Score 691.2; DB 14; Length 1487;
Best Local Similarity 80.1%; Pred. No. 3e-181;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGCTGAGATGTTGTAGACAGCACTGATATCGTGGGGATGCGAGCTTGCTGGAAA 1454
Db 12 GAGGGTGGGAGGAGCAGTACAGCACTGTATTGTGGGGGTGGCAACTCGCATGAAA 71
QY 1455 TGGTCAGAGAAGAGGTGAGATGCTAGAAAGGAAGGTGTTTCAAAAGAGTCTACTTG 1514
Db 72 TGGTCGAGCGCAGACAGCGAGAGTGCAGAAAGGAAGGAGGCTTCAAAAGATCTACTTG 131
QY 1515 CACCAAGAGGGAGTCTCTGGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGGTGC 1574
Db 132 CACCAAGAGGGGGTGGCCGACTCAAGAGGGGCTGTGTTTCACTTCTGTTGGGGT 191
QY 1575 GATG---TTTGTAGGGTATGAGTTGTACATGCTGCTGCTTTAGTAAGTCAAGCA 1631
Db 192 GATGCCACGCAAGGGGAGTGGTTATACATGCTGCTGCTTTGTAAGCCACTGGCT 251
QY 1632 CTTTCTCAAGGGTCTTGTCTGAACCAAGCATGTCAAGTCT---GCCATGTTCAACCA 1688
Db 252 CTTTACTCCAAGGATCTTATGAGAGCGTATGGCGCGCTCCAGCCATGATTCACCA 311
QY 1689 TCTGAGTAGCTGCCAAGGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAAGCGCT 1748
Db 312 TTGAGAGGAGCTCCCAAAGGTTCAATCTGAAAAGATCTGTTGAACCTGTGTAGGCGCT 371
QY 1749 GCCCTGTTAGTCGGTGTGGAATTCAGATCCTTCAACAGTTGCTGGAATTAACGGTGT 1808
Db 372 GCATTGTTGCTGGGTGTGGAATTCAGATGCTTCAAGCAGTTGCTGGAATTAATGAGTT 431
QY 1809 CTGTACTATAATCCCAAAATTCCTTGAGCAAGCTGTGTGGCAGTTATCTTTCCAAATTT 1868
Db 432 CTCTACTATACTCTCAAAATTCGAGCAAGCTGTGTGGCTGTCTTTCCAAATCTT 491
QY 1869 GGTCAGCTCGGCATCAGCATCATCTGTATCAGTTCTCTCACTACCTTACTAATGCTT 1928

Db 492 GGCCTCAGTTACAGATCAGCATCCATCTTGAATCAGTCTCTCACACCACTTACTCATGCTC 551
QY 1929 CCTTCATTTGGCTTGGCCATGCTGTATGATCTTCCGAGAGAAGTTTGTGCTGCTA 1988
Db 552 CCAAGCATTTGGTATGCCATGAGACTTATGATATATCTGGAAGAAGGTTTCTGCTACTG 611
QY 1989 GGCACAATTCATCTTGATAGCATCTCTAGTTATCTCTGTTGTCTCAATCTAATTGAT 2048
Db 612 GGCACAATTCATCTTGATAGCATCTCTAATTTGTTGGGTGTGTTCAATGTTATCAAC 671
QY 2049 TTGGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCTACTTCTGCTGC 2108
Db 672 TTGAGTACGGTGGCCCAAGCTGTGCTCTCCACAGTTAGCGTCATTTGTTCTGCTGC 731
QY 2109 TTCGTTATGGATTGGTCCCATCCCAACATTTATGTGAGAGATCTTCCAAACAGG 2168
Db 732 TTGTCATGGGCTTTGGCCCGATCCCAACATCTATGTGAGAGATTTCCCAACAGA 791
QY 2169 GTTCGTGCTCTGATTTGCCATTTGTGCTTTACATCTGTGATCGAGATATCATCGTC 2228
Db 792 GTCCGTGTGTCTGATCTGCTATTTGCGCCCTTCAATCTGTGATTTGTGACATTAATGTT 851
QY 2229 ACCTACAGCCTTCTGCTGATGCTGAATGCTATTTGAGACTGCGGGGTGTTTCAGCATATAT 2288
Db 852 ACCTACAGCCTGCTGATGCTGAATGCTATTTGCTAGCGGGTGTCTGTATATAT 911
QY 2289 GCACTGATGCTGATTTCTCTTGTGTTGCTCTTAAAGTCCCTGAGCAAAAGGGG 2348
Db 912 GCACTGCTTGTCTGATGCTCTTGTGTGCTTGTGCTTAAAGTCCCAAGAGCAAAAGGGC 971
QY 2349 ATGCCCTTGAAGTTATACCGAATCTTTGAGTGTGCGAAGCAAGC---GGCTGCA 2405
Db 972 ATGCCCTGAGGTCATCACCGAGTCTTTGCGGTGGGCGAAGCAAGCGCAGGCCACC 1031
QY 2406 AAAGCCTAATTTCTTGTACCTTTGTGTGCACTATTCACCTG 2449
Db 1032 ATTGCCTGATTCATCATGAGACTTTGTTTTCAGTTTGCACACTG 1075

RESULT 10

US-10-051-909-13
; Sequence 13, Application US/10051909
; Publication No. US2002019217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-909-13

Query Match 24.9%; Score 691.2; DB 14; Length 1487;
Best Local Similarity 80.1%; Pred. No. 3e-181;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGCTGAGATGTTGTAGACAGCACTGATATCGTGGGGATGCGAGCTTGCTGGAAA 1454
Db 12 GAGGGTGGGAGGAGCAGTACAGCACTGTATTTGGGGGTGGCAACTCGCATGAAA 71
QY 1455 TGGTCAGAGAAGAGTGAAGATGTAGAAAGGAAGGTGTTTCAAAAGAGTCTACTTG 1514

Db 72 TGGTCGGAGCGACAAAGCGAGGATGGCAAGAAAGAGGCTTCAAAAGAAATTAATTG 131
QY 1515 CACCAAGAGGAGTTCCTGGCTCAAGAAAGGGCTCAATTGTTCACTTCCCGGTGGC 1574
Db 132 CACCAAGAGGAGTTCCTGGCTCAAGAAAGGGCTCAATTGTTCACTTCCCGGTGGC 191
QY 1575 GATG---TTTGGAGGAGTGGAGTTGTATCATGCTGCTGCTTAAAGTCAAGTCA 1631
Db 192 GATGCCACGCAAGGGGAGTGGAGTTTATACATGCTGCTGCTTAAAGTCAAGTCA 251
QY 1632 CTTTCTCAAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1688
Db 252 CTTTCTCAAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 311
QY 1689 TCTGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1748
Db 312 TTGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 371
QY 1749 GCGCTGTAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1808
Db 372 GCAATGTTGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 431
QY 1809 CTGTACTATACCCCAAAATCTGAGCAAGCTGTTGAGCAATTTCTTCCAAATTT 1868
Db 432 CTCTACTATACCTCAAAATCTGAGCAAGCTGTTGAGCAATTTCTTCCAAATTT 491
QY 1869 GGTCTAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1928
Db 492 GCGCTCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 551
QY 1929 CCTGTGATGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1988
Db 552 CCAAGCATGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 611
QY 1989 GGCACAAATTCCTGATGATGATCTCTAGTATCTCTGCTGCTGCTGCTGCTGCTG 2048
Db 612 GGCACAAATTCCTGATGATGATCTCTAGTATCTCTGCTGCTGCTGCTGCTGCTG 671
QY 2049 TTGGGTACATGAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2108
Db 672 TTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 731
QY 2109 TTGCTATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2168
Db 732 TTGCTATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 791
QY 2169 GTTCGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2228
Db 792 GTTCGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 851
QY 2229 ACCTACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2288
Db 852 ACCTACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 911
QY 2289 GCAGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2348
Db 912 GCAGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 971
QY 2349 ATGCCCCCTGAGGTATATACCAATTTCTGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2405
Db 972 ATGCCCCCTGAGGTATATACCAATTTCTGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1031
QY 2406 AAAGCTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2449
Db 1032 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1075

RESULT 11
US-10-051-902-9
; Sequence 9, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-902-9

Query Match 21.4%; Score 594.6; DB 14; Length 1692;
Best Local Similarity 66.1%; Pred. No. 2.6e-154;
Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;

QY 929 CCGTATAGGAAACAAATCACACTTATGAGGCTGAGAGGCGCATGATGCTCG 988
Db 17 CAGAGAAAAGATCAAAATTAAGTTGATGAGCAAGAGGCGCATGCTGAGTCTAG 76
QY 989 ACCTCTAAGGAGCCATCATGCTTGAAGTGTCTTCTTGTGATCTGTCATGGAG 1048
Db 77 ACCTGTGCTGAGCAAAATCTGTGGC-----CTGTATCTAGAGAAAGAG 124
QY 1049 CATGTGAACAGAGTGTACCCCTTATGATCCGATTTGATGACACTTTTGTAGTCCA 1108
Db 125 CATGCAAAATCCAGCAGT---CTAGTGAACCTCTAGTGAACCTCTTGTAGTACA 181
QY 1109 TGAGATATGCTTCAAGCTGAGAGAGTATGAGAGACACATTTTCCAACTTTGGAAG 1168
Db 182 TGAGAGCTCCAGAAACAGAG-----AGCACCTTTTCCACACTTTGGAG 229
QY 1169 TATGTCAGTGTACAGATCAGATGCGCAAAATGAGCAGTGGAGTGAAGAATCTTCA 1228
Db 230 TATGTCAGTGTGGGGGAAATCAGCCAGAGATGAAGATTTGGAGTGAAGCCTAGC 289
QY 1229 TAGGATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1288
Db 290 CAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343
QY 1289 TAGCCATGCTGTCCAGGAGGCAAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 1348
Db 344 GAGTCATGATCTCAGTCAGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 400
QY 1349 TCACCGTGAAGTCTTTGAGCATGAGAGCAAACTCTTGAAGGAGGAGGAGGAGTGG 1408
Db 401 CCATAGTAACCTTGAAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 460
QY 1409 TGTGAGCAGCATGATATCGTGGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1468
Db 461 CACTGATGATGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 520
QY 1469 AGGTGAGATGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1528
Db 521 GGGCCAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
QY 1529 TCCTGGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1588
Db 581 TTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 637
QY 1589 TAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1648
Db 638 CAGTGAAGTGTATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697
QY 1649 TGCTGAACGAGCATGTCAGATGCTGCGCATGCTTACCCATCTGAGGTAGCTGCAAGG 1708
Db 698 TATGCTCAACGCGCAGTGTGAGCAGCATGATGATGATGATGATGATGATGATGATG 757

```
QY 1709 TTCACGTTGGAAGATTGTTGAACCTGAGTGAGCGGCTGCTGTAGTGGGTGG 1768
Db 758 GCCAAGTTGAGTATCTTTTGAACCTGGGTGAAGCATGATGATGTTGGGGTGG 817
QY 1769 AATTCAGATCCTTCAACAGTTTGTGGAATAAACGGTGTCTGTACTATACCCCAAAAT 1828
Db 818 AATGCAATTTCTTACAGCAGTTCTCTGTATATAATGGGTCTCTACTATACGCTCAAAAT 877
QY 1829 TCTTGAGCAAGCTGTGTGGAGTTATTTCTTCCAAATTTGGTCTAGCTGGGCTAGC 1888
Db 878 TCTTGAGCAAGCAGGTGTGGTTATCTTCTTCAAGCCTAGGCTTGGTCTACTTCTTC 937
QY 1889 ATCCATCTGATCAGTTCTCTCACTACCTTACTAATGCTTCTGCAATGGCTTTGCCAT 1948
Db 938 ATCTTCTTATAGTGGGTGAACAACCTGTGTAGTCTTCTGTATAGCAATGGCCAT 997
QY 1949 GCTGCTTATGATCTTTCGGAAGAAGGTTTGTGCTGAGGACAAATTCATCTTGAT 2008
Db 998 GAGGCTCATGATATTTCAAGCAGAAGAGACTTGTGCTCAGTACAATCCCGTCTAAT 1057
QY 2009 AGCATCTAGTATCTGTGTGTGTCCATCTAATGATTTGGGTACACTAGCCCATGC 2068
Db 1058 AGCAGCTCTCTCATATATAGTCTCTGGAAGTCTGTGATTTGGGATCCACTGCAAAATGC 1117
QY 2069 TTTGCTCTCCACCGTCAAGTGTATCTGTCTACTTCTGTCTCTGCTTCTGATGGATTTGCC 2128
Db 1118 ATCAATCTCAACCATTAAGTGTATGTTCTATTTCTGTCTTCTTGTCTGATGGATTTGCAAC 1177
QY 2129 CATCCCCAATTTTATGTGAGAGATCTTTCACCAAGGTTGCTGCTCTGATTTGC 2188
Db 1178 AATTCCTAATATACTTTGTGAGAGATCTTCCCATGAGTTCGTGCTCTGCAATTC 1237
QY 2189 CATTTGCTTACATCTGTGATCGAGATATCATGCTCACTACAGCCTTCTGTGAT 2248
Db 1238 TATTTGCTTACCTTGTGATCTGTGATATCATTTGTCACCTACACACTCCAGTTAT 1297
QY 2249 GCTGAATGCTATGTGAGTGGCGGTGTTTTCAGCATATATGACATGCTATGCTGATTC 2308
Db 1298 GCTCAATCTGATAGGCTCGGTGTGTTTGTGATTTATGCTGTGTGCTTCAATAGC 1357
QY 2309 CTTTGTGTCTTCTTCTTAAGTCTCCCTGAGACAAAGGGATGCCCCCTGAGTTATAC 2368
Db 1358 ATGGGTGTTGCTTTTGAAGTTCAGAAAACCAAGGCGATGCCACTGGAATGATCAT 1417
QY 2369 CGAATCTTTGAGTGTGCGAAGCA 2395
Db 1418 TGAGTTCTTCTGTGCGAGCAAAACA 1444
```

RESULT 12

```
US-10-051-909-9
; Sequence 9, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-909-9
```

```
Query Match 21.4%; Score 594.6; DB 14; Length 1692;
Best Local Similarity 66.1%; Pred. No. 2.6e-154;
Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;
QY 929 CGGTATAGGACAAATCAACACTTTATGGGCTGAAGAGGCCAGTCAATGCTCG 988
Db 17 CAGAGAAAAGATCAAAATTAAGTTGATGACCAAGAACAGGCCAGTCTGGGTGCTAG 76
QY 989 ACCCTTAAGGACCCATCATGCTTGGAGTGTGCTTCTTGTGATCTGTATGGAG 1048
Db 77 ACCTGTGCTGACCAAAATCTGTGGC-----CTGTATCAAGAAAGAG 124
QY 1049 CATGTGAACCAAGTGTACCCCTTATGATCCGATTTGTGACACTTTTGTAGTGTCCA 1108
Db 125 CATGCAAAATCCAGCAGT---CTAGTGACCTCTAGTGACCCCTTGTGTAGTGTACA 181
QY 1109 TGAGATATGCTCAAGCTGAGGAGATATGAGACACATTTTCCAACTTTGGAAG 1168
Db 182 TGAGAGCTCCCAAGAACAGGA-----AGCACCTTTTCCACACTTTGGAG 229
QY 1169 TATGTTAGTGTACAGATCAGCATGCCAAATAGACAGTGGGATGAGAAATCTTCA 1228
Db 230 TATGTTAGTGTGGGGGAAATCAGCCAAAGATGAAGATTGGATGAGAAAGCCTAGC 289
QY 1229 TAGGATGACGAGAGTACGATCTGATGTGACAGAGTGAATGAGCAAAATCTCCA 1288
Db 290 CAGAGAGGTGATGATATATGCTCTGATGCT-----GGTATCTGATGACAAATTTGCA 343
QY 1289 TAGCCATGCTGTCCAGAGCAGCAAGTGTGGAAGGAGGAGGAGGAGGAGGAGG 1348
Db 344 GAGTCCATGATCTCAGTCAACCAAGAGCTGGA--TAAGACATACCTCCTCATGC 400
QY 1349 TCACGTTGAGTCTTTGAGCATGAGAGGCAAAACCTCTTAGGGAGGGTGAAGATGG 1408
Db 401 CCATAGTAACCTTGCAAGCATGAGGCAAGATGATCTTTACATGGAATTCAGAGAAACC 460
QY 1409 TGTGAGCAGCATGATATCGGTGGGGAGTGGCAGCTTGTGGAATGTTCAGAGAAAGA 1468
Db 461 CACTGATGATGAGTGGATTTGTGTGTGGCAGCTAGCATGGAATGTCTGAAAAGAGA 520
QY 1469 AGGTGAGATGTGAAGAGAGAGTGTTCAAAGAGTCTACTTGCACCAAGAGGAGT 1528
Db 521 GGGCCAGATGAGAAAGAGAGAGTGTTCAAAGATATATTACCAAGATGTGG 580
QY 1529 TCCTGCTCAAGAGGGCTCAATTTGTTCACTCCGGTGTGGCGATGTTTGGAGG 1588
Db 581 TTCTGATCTAGACGTGGGTCTGTGTTCACTCCC--TGCGGTGATTTTCAACTGA 637
QY 1589 TAGTAGTTGTACATGCTGCTGCTTTAGTAACTAGTACAGCACTTTTCAAGGGTCT 1648
Db 638 CAGTGAAGTTGTACAGGCTGCTGCTGTGTGATGATCAGCCTGCCCTTATTAATGAGACCT 697
QY 1649 TGCTGAACCAAGCATGTACAGATGCTGCCATGTTCAACCATCTGAGGTAGCTGCCAAGG 1708
Db 698 TATGCTCAACGGCCAGTTGAGCAGCTATGATTCATCCCTGTAACAAATTGCCAAGAG 757
QY 1709 TTCACGTTGGAAGATTGTTTGAACCTGAGTGAAGGCGTGTCTGTAGTGGGTGG 1768
Db 758 GCCAAGTTGAGTATCTTTTGAACCTGGGTGAAGCATGATGATGTTGGGGTGG 817
QY 1769 AATTCAGATCCTTCAACAGTTTGTGGAATAAACGGTGTCTGTACTATACCCCAAAAT 1828
Db 818 AATGCAATTTCTTACAGCAGTTCTCTGTATATAATGGGTCTCTACTATACGCTCAAAAT 877
QY 1829 TCTTGAGCAAGCTGTGTGGAGTTATTTCTTCCAAATTTGGTCTAGCTGGGATCAGC 1888
Db 878 TCTTGAGCAAGCAGGTGTGGTTATCTTCTTCAAGCCTAGGCTTGTCTACTTCTTC 937
QY 1889 ATCCATCTGATCAGTTCTCTCACTACCTTAAATGCTTCTGATGGCTTTGCCAT 1948
Db 938 ATCTTCTTATAGTGGGTGAACAACCTGTGTAGTCTTCTGTATAGCAATGGCCAT 997
```


QY	1949	GCTGCTTATGATCTTTCCGAGAAGGTTTTTGCTGCTAGGCACAATTCCAATCTGAT	2008
Db	998	GAGGCTCATGATATTTCAAGCAGAAGGACTTTGCTGCTCAGTACAATCCCCGCTTAAT	1057
QY	2009	AGCATCTCTAGTTATCCGTGTGTGTCCAAATCTAAATTGAATTGGGTACACTAGCCCATGC	2068
Db	1058	AGCAGCTCTTCTCATATATATAGTCTGGGAAGTCTTGGAATTGGGATCCACTGCAATGC	1117
QY	2069	TTTGCTCTCCACCGTCAGTGTATATGCTCTACTTCTGCTGCTTGTATGGGATTTGGTCC	2128
Db	1118	ATCAATCTCAACCATATAGTGTATATGTCTATTTCTGTTCTTTGTGCATGGGATTTGGACC	1177
QY	2129	CATCCCAACATTTTATATGTGCAGAGATCTTCCAAACAGGGTTGGCTCTGTATTC	2188
Db	1178	AATTCCTAATATACTTTGTGCAGAGATCTTCCCACTCGAGTTGGTCTCTGCATTTGC	1237
QY	2189	CATTTGTGCTTACATTTCTGGATCCGAGATATCATCTCAACCTACAGCCTTCTGTGAT	2248
Db	1238	TATTTGTGCCCTTACCCTTTTGATCTGTGATATCATTTGCACCTACACACTCCCAAGTAT	1297
QY	2249	GCTGAATGCTATTTGACCTGGCGGGTGTTCAGCAATATATGCAGTCTGATGCTTGAATTC	2308
Db	1298	GCTCAATCTGTAGGCGCTCGCTGTGTTTTTGGTATTTATGCTGCTGTGCTTCAATAGC	1357
QY	2309	CTTGTGTGCTCTTCCCTTAAGTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTAC	2368
Db	1358	ATGGGTGTTGTCTTTTGAAGAATTCAGAAAACCAAGGGCATGCCACTGGAAGTGAATCAT	1417
QY	2369	CGAATCTTTTGACAGTTGGTGGGAAGCA	2395
Db	1418	TGAGTCTTCTCTGTGCGAGCAAAACA	1444

```

RESULT 13
US-10-051-902-15
; Sequence 15, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-15

```

Query Match	14.2%;	Score 394.4;	DE 14;	Length 1009;
Best Local Similarity	74.4%;	Pred. No. 1.1e-98;		
Matches 497;	Conservative	0;	Mismatches 171;	Indels 0;
			Gaps	0;

Oy	1730	TGAACCTGGAGTGAAGCCGTGCCCTGTTAGTCGGTGTTGGAAATTCAATTCAGATCCTTCAACAGTT	1789
Db	1	TGAACCTGGAGTGAAGCATGCACACTGTTGTTGGTAAGAATTACAGATCCTGCACAGATT	60
Oy	1790	TGCTGGAATPAACCGTGTTCTGTACTATAACCCACAATTTCTTGAGCAAGCTGGTGGC	1849
Db	61	TGCGGGTATCAATGAGAGTCTCTACTACACACACTCAGATACCTTGAGCAAGCAGGTGTGG	120
Oy	1850	AGTTATTTCTTCCAAATTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAATTCTCT	1909
Db	121	GGTTCCTTCTATCAACAATGGAAGCTCTTCCTCAGCATCTATCTTATTATGTCCTT	180
Oy	1910	CACCTACCCTTACTAATGCTTCTCTGCATTGGCTTTGCCATGCTGCTTATGAGATCTTCCGG	1969

Db	181	GACAACTTGCTGATGCTTCCCAAGCATGGCATGCCCATGAGACTCATGATATGTACAGG	240
QY	1970	AAGAAGGTTTTGCTGCTAGGCAACAATTCCAAATCTTGATAGCATCTAGTTATCTGTG	20239
Db	241	AAGAAGGTTCTCTCTCTTCAACAATCCCTGTCTTGATAGTAGCGTAGCTGTCTGTG	300
QY	2030	TGTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGT	20899
Db	301	TTTAGTAATGTTCTGGATGTGGAACCATGTGACGCTGCGCTCTCAACGATCAGCGT	360
QY	2090	TATCGTCTACTTCTGCTGCTTGTGTTATGGATTTGGTCCCATCCCAACATTTATGTGC	21499
Db	361	CATCGTCTATTTCTGCTTCTTGTGTCATGGGTTGGGCCATCCCAATATTTCTGTGCGC	420
QY	2150	AGAGATCTTTCCCAACCAAGGTTGCTGGCCTCTGTATGCCATTTGTGCCTTACATTTG	22099
Db	421	GGAGATTTTCCCACTCTGTCTCGTGCGATCTGCATAGCCATCTGCGCGCTAACCTTCTG	480
QY	2210	GATCGAGATATCATCTGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTGCACTGGC	22699
Db	481	GATCGCGACATCATCTGACATACACTCTCCCGTGATGCTCAATGCCAATGGTCTCGC	540
QY	2270	GGGTGTTTTACGATATATGCAAGTCGATGCTGATTTCTTTGTGTGCTTCTCTTA	23299
Db	541	TGGAGTCTTCCGCATATATGCCATCGTTTGTGTACTAGCCTTTGTATTCGTACATGAA	600
QY	2330	GCTCCCTGAGACAAAGGGGATGCCCCCTGAGGTTATTAACGAATTTCTTGCAAGTGTGC	23899
Db	601	GCTCCCTGAGACAAAGGGCATGCCCCCTGAGGTCATCACCGAGTTCCTCTGTGTGCGGC	660
QY	2390	GAAGCAAG 2397	
Db	661	AAAGCAGG 668	

```

RESULT 14
US-10-051-909-15
; Sequence 15, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B0163 US ClP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Trilicium aestivum
US-10-051-909-15

```

Query Match	14.2%;	Score 394.4;	DB 14;	Length 1009;
Best Local Similarity	74.4%;	Pred. No. 1.1e-98;		
Matches 497; Conservative	0;	Mismatches 171;	Indels 0;	Gaps 0;

[illegible]

```
Db 121 GGTCTCTATCAAAACATGGACTAAGCTCTTCCTCAGCATCTATTCTATTAGCCCTT 180
QY 1910 CACTACCTTACTAATGCTTCCTGCATTTGGCTTGGCCATGCTGCTTATGATCTTCCGG 1969
Db 181 GACAACCTTGCTGATGCTTCCCAAGCATTTGGCATCGCCATGAGACTCATGGATATGTCAAG 240
QY 1970 AAGAAGTTTGGCTGCTAGGACAAATTCCAATCTTGATAGCATCTCTAGTATCTTGGT 2029
Db 241 AAGAAGTTTCTCTCTCTTCAACAATCCCTGCTTGATAGTAGCGCTAGCTGTCTGGT 300
QY 2030 TGTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCACCGTCAAGTGT 2089
Db 301 TTTAGTGAATGTTCTGGAATGTCGGAACCATGATGACGCTGCGCTCTCAACGATCAGCGT 360
QY 2090 TATCGTCTACTTCTGCTGCTTGTATGGATTTGGTCCCATCCCAACATTTATGTGC 2149
Db 361 CATCGTCTATTTCTGCTTCTGTCATGGGTTGGGCTATCCCAATATTTCTGCGC 420
QY 2150 AGAGATCTTCCCAACGAGGTTGCGCTGCTGTATGCCATTTGTGCTTACATTTCTG 2209
Db 421 GGAGATTTTCCCAACCTCTGTCGCTGCGCATCTGCATAGCCATCTGCGCGTACCTTCTG 480
QY 2210 GATCGGAGATATCATCTGTCACCTACAGCCCTCTGTGATGCTGAATGCTATTGGACTGGC 2269
Db 481 GATCGGCGACATCATCTGTCATACACTCTCCCGTGTGCTCAATGCCAATGGTCTCGC 540
QY 2270 GGGTGTTCAGCATATATGAGTGTGATGCTGATTTCTTGTGCTTCTTCTTAA 2329
Db 541 TGAAGTCTTCGGCATATATGCCATCGTTGTGTACTAGCCTTTGTATTGCTTACATGAA 600
QY 2330 GGTCCCTGAGACAAAGGGGATGCCCCCTTGAGGTTATTAACGAATTTTGCAGTTGGTGC 2389
Db 601 GTTCCCTGAGACAAAGGGGATGCCCCCTTGAGGTCATACCGAGTTCTTCTGTGCGGGC 660
QY 2390 GAAGCAAG 2397
Db 661 AAAGCAGG 668
```

RESULT 15

```
US-09-923-876-6145
; Sequence 6145, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 6145
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700458459H1
; NAME/KEY: unsure
; LOCATION: 3, 16, 20
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-6145
```

```
Query Match 9.7%; Score 269.2; DB 9; Length 285;
Best Local Similarity 97.8%; Pred. No. 3.2e-64;
Matches 271; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1433 GGGATGGCAGCTTGCTTGAATGTCAGAGAAGAGTGAGAAATGGTAGAAGAAGG 1492
Db 1 GGNATGGCAGCTTGCTTGNATGTCAGAGAAGAGTGAGAAATGGTAGAAGAAGG 60
QY 1493 TGGTTCAAAAGAGTCTACTTGCAACCAAGAGGAGTTCCTGGCTCAAGAGGGGCTCAAT 1552
Db 61 TGGTTCAAAAGAGTCTACTTGCAACCAAGAGGAGTTCCTGGCTCAAGAGGGGCTCAAT 120
QY 1553 TGTTCACCTTCCCGGTGCTGCGGATGTTTGAAGGATGAGTGTGTACATGCTGCTGC 1612
Db 121 TGTTCACCTTCCCGGTGCTGCGGATGTTTGAAGGATGAGTGTGTACATGCTGCTGC 180
QY 1613 TTTAGTAACTCAGTCAGCACTTTCTCAAGAGGCTTGTGAACCAACGATGTCAGATGC 1672
Db 181 TTTAGTAACTCAGTCAGCACTTTCTCAAGAGGCTTGTGAACCAACGATGTCAGATGC 240
QY 1673 TGCCATGTTTCAACCATCTGAGTAGCTGCCAAAGT 1709
Db 241 TGCCATGTTTCAACCATCTGAGTAGCTGCCAAAGT 277
```

Search completed: January 5, 2004, 18:53:59
Job time : 885 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 09:53:40 ; Search time 5656 Seconds
(without alignments)
11933.085 Million cell updates/sec

Title: US-10-051-909-31

Perfect score: 2777
Sequence: 1 caccgggttagatcgagc.....aaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2126.2	76.6	2180	11	AY105508 Zea mays
2	810.4	29.2	911	29	CC423708 PUFOR49TB
3	799.8	28.8	845	29	BZ723844 PUCF160TD
4	732	26.4	795	29	BZ989776 PUGDY07TD

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
5	594	21.4	594	9	AI861088							
6	553.2	19.9	817	29	BZ819234							
7	531	19.1	829	13	BQ752737							
8	525	18.9	545	12	BI992011							
9	508.6	18.3	572	12	BM325261							
10	491.8	17.7	697	14	CA486683							
11	469.6	16.9	664	12	BI074258							
12	466.2	16.8	653	10	BG356047							
13	465	16.7	661	12	BM329724							
14	464.6	16.7	471	29	CC004220							
15	460	16.6	703	14	CA502490							
16	456.2	16.4	494	12	BM378346							
17	451.6	16.3	624	14	CB213154							
18	448.4	16.1	647	10	BG102713							
19	443.8	16.0	656	13	BU925675							
20	437	15.7	634	10	BG411950							
21	432	15.6	618	10	BG605795							
22	424.6	15.3	623	14	CA619074							
23	420.2	15.1	489	10	BG103044							
24	415.2	15.0	620	13	BU989641							
25	411.6	14.8	596	14	CB924656							
26	406.8	14.6	645	14	CA030046							
27	406.6	14.6	579	10	BG103781							
28	405.8	14.6	688	14	CA502649							
29	403.6	14.5	645	13	BQ240484							
30	398.8	14.4	451	10	BG649604							
31	386.4	13.9	745	14	CB635485							
32	386	13.9	552	10	BG355585							
33	380.8	13.7	604	9	AV916059							
34	379.2	13.7	879	29	CC423712							
35	378	13.6	847	14	CB981615							
36	376.2	13.5	550	12	BM374503							
37	376.2	13.5	762	12	BI310999							
38	375.2	13.5	601	13	BQ246792							
39	375	13.5	557	12	BI098058							
40	372.6	13.4	749	13	BU040051							
41	370.8	13.4	671	9	AV909064							
42	368.2	13.3	859	12	BM817360							
43	367.2	13.2	617	13	BU925683							
44	363.2	13.1	579	13	BU925708							
45	361.8	13.0	653	12	BJ450590							

ALIGNMENTS

RESULT 1
AY105508
LOCUS
Zea mays PC011453 mRNA sequence.
DEFINITION
AY105508
VERSION
AY105508.1 GI:21208586
KEYWORDS
HTC.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

www.zimbardo.iaestate.edu.

FEATURES	Location/Qualifiers
source	1. .2180

```
/organism="Zea mays"  
/mol_type="mRNA"  
/db_xref="MaizeDB:636667"  
/db_xref="taxon:4577"  
/clone_lib="Maize Mapping Project/DuPont Cornsensus  
Library"
```

/note="this sequence is part of a project of ESR assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT	533 a	447 c	555 g	645 t
ORIGIN				

Query Match	76.6%;	Score 2126.2;	DB 11;	Length 2180;
Best Local Similarity	99.6%;	Pred. No. 5.2e-242;		
Matches 2131;	Conservative	0;	Mismatches 8;	Indels 0;
				Gaps 0;

QY	619	CATACTGCATGGTGTGGGATGTCCTCTGCCCATCACCCGATTGGAGAAATTATGCTTG	678
Db	42	CGCTTTTCATGNGTGGGATGTCCTGTCGCCATCACCCGATTGGAGAAATTATGCTTG	101
QY	679	GTCGTCTCGCGATACCTTCATGTGTTCTTTGGTTGACAATAATTATCTTCCGTAAT	738
Db	102	GTCGTCTCGCGATACCTTCATGTGTTCTTTGGTTGACAATAATTATCTTCCGTAAT	161
QY	739	CTCCAAGATGGCTCGTTAGCAAAAGTCGGATGGCAGAGGCCAAAAAGGTGTGCAAAAGT	798
Db	162	CTCCAAGATGGCTCGTTAGCAAAAGTCGGATGGCAGAGGCCAAAAAGGTGTGCAAAAGT	221
QY	799	TACGGGGGAAAAGACGATGTCAGGTGAATTGTCCTTCTTCGAAAGGCTTGAGGTTG	858
Db	222	TACGGTGGAAAGACGATGTCAGGTGAATTGTCCTTCTTCGAAAGGCTTGAGGTTG	281
QY	859	GAGGAGACACTTCCATTGAAGATACATCATTTGGACCTGCCACCGAGCAGCCGATGATC	918
Db	282	GAGGAGACACTTCCATTGAAGATACATCATTTGGACCTGCCACCGAGCAGCCGATGATC	341
QY	919	TTGTTACTGACGGTGATGAAGAACAAATCACACTTATGGGCTTGAAAGGCCAGTCAT	978
Db	342	TTGTTACTGACGGTGATGAAGAACAAATCACACTTATGGGCTTGAAAGGCCAGTCAT	401
QY	979	GGATTGCTCGACCTTCTAAGGGACCCATCATGCTTGAAGTGTGCTTCTTGCATCTC	1038
Db	402	GGATTGCTCGACCTTCTAAGGGACCCATCATGCTTGAAGTGTGCTTCTTGCATCTC	461
QY	1039	GTCATGGGAGCATGGTGAACCCAGAGTGTACCCCTTATGATCCGATGTGACACTTTTGG	1098
Db	462	GTCATGGGAGCATGGTGAACCCAGAGTGTACCCCTTATGATCCGATGTGACACTTTTGG	521
QY	1099	GTAAGTCCATGAGAATATGCTCAAGCTGAGAGAAAGTATGAGGAGCACATTTGTTCCAA	1156
Db	522	GTAAGTCCATGAGAATATGCTCAAGCTGAGAGAAAGTATGAGGAGCACATTTGTTCCAA	581
QY	1159	ACTTTGGAAGTATGTTCAGTGTCAAGATCAGCATGCCAAAAATGACAGTGGGATGAAG	1218
Db	582	ACTTTGGAAGTATGTTCAGTGTCAAGATCAGCATGCCAAAAATGACAGTGGGATGAAG	641
QY	1219	AGAAATTTCAATAGGGATGACGAGAGTACCAATCTGATGTTGCAGAGGTGACTATGAGG	1278
Db	642	AGAAATTTCAATAGGGATGACGAGAGTACCAATCTGATGTTGCAGAGGTGACTATGAGG	701
QY	1279	ACAAATTCCTATAGCCCATTTGCTGTCCAGGCAAGCAACAGGTGCGGAAGGGAAGACATTG	1338
Db	702	ACAAATTCCTATAGCCCATTTGCTGTCCAGGCAAGCAACAGGTGCGGAAGGGAAGACATTG	761
QY	1339	TGACCATGCTCACCCGTGGAAGTGTGAGCATGAGAAAGCCAAACCTCTTAGGGGAGG	1398

Db	762	TGCACCATGGTCAACCGTGGAGAGTGTCTTGAGCATGAGAAGGCAAGCCTCTTAGGGGAGG	821
QY	1399	GTGAGATGGTGTAGCAGCACTGATATCCGGTGGGGGATGGCAGCTTGTGGAAATGGT	1458
Db	822	GTGAGATGGTGTAGCAGCACTGATATCCGGTGGGGGATGGCAGCTTGTGGAAATGGT	881
QY	1459	CAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGTTCAAAAGAGTCTACTTGCACC	1518
Db	882	CAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGTTCAAAAGAGTCTACTTGCACC	941
QY	1519	AAGAGGAGTTCCTGGCTCAAGAAGGGGCTCAATTGTTCACTTCCCGGTGTGGCGATG	1578
Db	942	AAGAGGAGTTCCTGGCTCAAGAAGGGGCTCAATTGTTCACTTCCCGGTGTGGCGATG	1001
QY	1579	TTTTTGGGGTAGTGAATTGTATACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTTTTCT	1638
Db	1002	TTTTTGGGGTAGTGAATTGTATACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTTTTCT	1061
QY	1639	CAAAGGCTTGTCTGAACCAAGCATGTACAGTGTCCATGGTTCACCCATCTGAGTAG	1698
Db	1062	CAAAGGCTTGTCTGAACCAAGCATGTACAGTGTCCATGGTTCACCCATCTGAGTAG	1121
QY	1699	CTGCCAAAGGTTCAACGTTGGAAAGATTTGTTTGAACCTGGAGTGAGGCGTCCCTGTAG	1758
Db	1122	CTGCCAAAGGTTCAACGTTGGAAAGATTTGTTTGAACCTGGAGTGAGGCGTCCCTGTAG	1181
QY	1759	TCGGTGTGGAAATTCAGATCTTCAACAGTTTGTCTGGAATAAACGGTGTCTGTACTATA	1818
Db	1182	TCGGTGTGGAAATTCAGATCTTCAACAGTTTGTCTGGAATAAACGGTGTCTGTACTATA	1241
QY	1819	CCCCCAAAATTCCTTGAGCAAGCTGGTGTGGCAGTTATCTTCCAAATTTGGTCTCAGCT	1878
Db	1242	CCCCCAAAATTCCTTGAGCAAGCTGGTGTGGCAGTTATCTTCCAAATTTGGTCTCAGCT	1301
QY	1879	CGGCATCAGCATCCATCTTGATCAGTTCTCTCACTAACCCTAATGCTTCCCTGCAATG	1938
Db	1302	CGGCATCAGCATCCATCTTGATCAGTTCTCTCACTAACCCTAATGCTTCCCTGCAATG	1361
QY	1939	GCTTTGCCATGCTGCTTATGGAATCTTCCGAAGAAGGTTTTTGTCTGTAGGCACAATTC	1998
Db	1362	GCTTTGCCATGCTGCTTATGGAATCTTCCGAAGAAGGTTTTTGTCTGTAGGCACAATTC	1421
QY	1999	CAATCTTGATAGCATCTCTAGTTATCTCTGGTGTGTCCAAATCTAATGATTTGGGTACAC	2058
Db	1422	CAATCTTGATAGCATCTCTAGTTATCTCTGGTGTGTCCAAATCTAATGATTTGGGTACAC	1481
QY	2059	TAGCCCATGCTTTGCTCTCCACCGTCAGTGTATGCTTACTTCTGCTGCTTCGTTATGG	2118
Db	1482	TAGCCCATGCTTTGCTCTCCACCGTCAGTGTATGCTTACTTCTGCTGCTTCGTTATGG	1541
QY	2119	GATTTGGTCCCATCCCAACATTTTATGTGCAGAGATCTTCCAAACAGGTTGTTGCC	2178
Db	1542	GATTTGGTCCCATCCCAACATTTTATGTGCAGAGATCTTCCAAACAGGTTGTTGCC	1601
QY	2179	TCTGATTTGCCATTTGTGCTTTACATTTCTGATCGGAGATATCATGTCACCTACAGCC	2238
Db	1602	TCTGATTTGCCATTTGTGCTTTACATTTCTGATCGGAGATATCATGTCACCTACAGCC	1661
QY	2239	TTTCTGTGATGCTGAATGCTAATTTGGAAGTGGCGGGTGTTCAGCATATATGCAATCGTAT	2298
Db	1662	TTTCTGTGATGCTGAATGCTAATTTGGAAGTGGCGGGTGTTCAGCATATATGCAATCGTAT	1721
QY	2299	GCTTGATTTCTTGTGTGCTCTTCTTAAGGTCCCTTGAGACAAGAGGGGATGCCCTTG	2358
Db	1722	GCTTGATTTCTTGTGTGCTCTTCTTAAGGTCCCTTGAGACAAGAGGGGATGCCCTTG	1781
QY	2359	AGGTTATTAACCGAATCTTTGCAGTGTGTGCGAAGCAAGCGGCTGCAAAAAGCCTAATTTT	2418
Db	1782	AGGTTATTAACCGAATCTTTGCAGTGTGTGCGAAGCAAGCGGCTGCAAAAAGCCTAATTTT	1841
QY	2419	TTTGTACCTTTGTGTGCACTAATTTGCACTGTAAAGTTAAGAACTTGAAGGGGTTTCAACA	2478

Db 1842 TTGGTACCTTTGNGTGAACATAATTGCACGTGTAAGTTAGAACTTGAAGGGGTTTCACCA 1901

QY 2479 AGAAGCTCGAGAATTAATTGGATTGTGTAAATGTTAAGGGAACGAACATCTGCTCAT 2538

Db 1902 AGAAGCTCGAGAATTAATTGGATTGTGTAAATGTTAAGGGAACGAACATCTGCTCAT 1961

QY 2539 GCTCCTCAAAACGGTAAAGAGTCCCTCAATGGCAATAGAGTGGTTAAAGTTGTCAT 2598

Db 1962 GCTCCTCAAAACGGTAAAGAGTCCCTCAATGGCAATAGAGTGGTTAAAGTTGTCAT 2021

QY 2599 GTCAATTACCATATGTTTACCTATTGTACTGTATTAAGTCAAGCTATTCAAGCTG 2658

Db 2022 GTCAATTACCATATGTTTACCTATTGTACTGTATTAAGTCAAGCTATTCAAGCTG 2081

QY 2659 GTTGTGCTAGAAATCTTTAGAACAAAGATGATATGATCTGATCTGATGTTAATATT 2718

Db 2082 GTTGTGCTAGAAATCTTTAGAACAAAGATGATATGATCTGATCTGATGTTAATATT 2141

QY 2719 CAAATCTCAAAATAAGAAATATCGTTTCTCAAAAAA 2757

Db 2142 CAAATCTCAAAATAAGAAATATCGTTTCTCAAAAAA 2180

RESULT 2
CC423708/c 911 bp DNA linear GSS 19-MAY-2003

LOCUS PUHOR49TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM5BTa49102,
DEFINITION genomic survey sequence.
ACCESSION CC423708
VERSION CC423708.1 GI:30903798
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 911)

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished

TITLE
JOURNAL Other_GSSs: PUHOR49TD
COMMENT Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR

Class: sheared ends.
FEATURES
Location/Qualifiers
1. 911

Source
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BTa49102"
/clone_11b="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

BASE COUNT 242 a 264 c 163 g 242 t

ORIGIN
Query Match 29.2%; Score 810.4; DB 29; Length 911;
Best Local Similarity 99.3%; Pred. No. 1.1e-86;
Matches 814; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 817 TCTCAGGTGAATTTGCTCTTCTCGAAGGTTGAGGTTGAGAGACACTTCATTG 876

Db 820 TCACAGGTGAATTTGCTCTTCTCGAAGGTTGAGGTTGAGAGACACTTCATTG 761

QY 877 AAGAGTACATCATTTGACCTGCCACCGAGCGCGATGATCTTGTACTGACGGTGATA 936

Db 760 AAGAGTACATCATTTGACCTGCCACCGAGCGCGATGATCTTGTACTGACGGTGATA 701

QY 937 AGAACAAATCACACTTATGGGCTTGAGAGAGCCAGTCATGATTTGCTCGACCTTCTA 996

Db 700 AGAACAAATCACACTTATGGGCTTGAGAGAGCCAGTCATGATTTGCTCGACCTTCTA 641

QY 997 AGGACCCCATCATGCTTGAAGTGTGCTTCTCTTGACATCTCGTCAATGGAGCATGTGA 1056

Db 640 AGGACCCCATCATGCTTGAAGTGTGCTTCTCTTGACATCTCGTCAATGGAGCATGTGA 581

QY 1057 ACCAGAGTACCCCTTATGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGATA 1116

Db 580 ACCAGAGTACCCCTTATGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGATA 521

QY 1117 TGCCTCAAGCTGAGAGGAAGTATGAGAGACATTGTTCCAACTTTGGAAGTATGTCA 1176

Db 520 TGCCTCAAGCTGAGAGGAAGTATGAGAGACATTGTTCCAACTTTGGAAGTATGTCA 461

QY 1177 GTGTACAGATCAGCATGCCAAAAATGACAGTGGAGTGAAGAAATCTTCATAGGATG 1236

Db 460 GTGTACAGATCAGCATGCCAAAAATGACAGTGGAGTGAAGAAATCTTCATAGGATG 401

QY 1237 ACAGAGTACGATCTGATGTTGTCAGAGGTGACTATGAGACAAATCTCCATAGCCCAT 1296

Db 400 ACAGAGTACGATCTGATGTTGTCAGAGGTGACTATGAGACAAATCTCCATAGCCCAT 341

QY 1297 TGCTGTCCAGGACGACACAGGTGCGGAAGGAAGACATTTGTGACCATGTCACCGTG 1356

Db 340 TGCTGTCCAGGACGACACAGGTGCGGAAGGAAGACATTTGTGACCATGTCACCGTG 281

QY 1357 GAAGTGTCTTGACATGAGAAAGGCAACCTCTTAAGGGAGGGTGAATGCTGTAGCA 1416

Db 280 GAAGTGTCTTGACATGAGAAAGGCAACCTCTTAAGGGAGGGTGAATGCTGTAGCA 221

QY 1417 GCAGTATATCGGTGGGGGATGGCAGCTTCTTGAAATGTTGAGAGAGAGGTGAGA 1476

Db 220 GCAGTATATCGGTGGGGGATGGCAGCTTCTTGAAATGTTGAGAGAGAGGTGAGA 161

QY 1477 ATGTAGAAAGAGGTGTTTCAAAAGAGTCTACTTGACCAAGAGGAGTTCTGCTG 1536

Db 160 ATGTAGAAAGAGGTGTTTCAAAAGAGTCTACTTGACCAAGAGGAGTTCTGCTGCT 101

QY 1537 CAAGAAGGGCTCAATTTTCACTTCCCGGTGTCGCGATGTTTTCAGCGTAGTAGT 1596

Db 100 CAAGAAGGGCTCAATTTTCACTTCCCGGTGTCGCGATGTTTTCAGCGTAGTAGT 41

QY 1597 TTGTACATGCTGCTGCTTTAGTAAGTCAGTCACACTTT 1636

Db 40 TTGTACATGCTGCTGCTTTAGTAAGTCAGTCACACTTT 1

RESULT 3
BZ723844 845 bp DNA linear GSS 24-FEB-2003

LOCUS PUCFI60TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM5BTa133J23,
DEFINITION genomic survey sequence.
ACCESSION BZ723844
VERSION BZ723844.1 GI:28519500
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 845)

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished

TITLE
JOURNAL Contact: Cathy Whitelaw
COMMENT TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

Fax: 301-838-0208
Email: white@lwtigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source location/Qualifiers
1..845
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM081A33J23"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

BASE COUNT 214 a 156 c 259 g 216 t

Query Match 28.8%; Score 799.8; DB 29; Length 845;
Best Local Similarity 98.8%; Pred. No. 2e-85;
Matches 837; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

QY 916 ATCTGTACTGACGGTGATAGAAACAATCACACTTTATGGGCTGAAGAGGCCAGT 975
Db 1 ATCTGTACTGACGGTGATAGAAACAATCACACTTTATGGGCTGAAGAGGCCAGT 60
QY 976 CATGATTTGCTGACCTTCTAAGGAGCCCATCATGCTTGAAGTGTCTTCTGAT 1035
Db 61 CATGATTTGCTGACCTTCTAAGGAGCCCATCATGCTTGAAGTGTCTTCTGAT 120
QY 1036 CTCGTATGGGAGCATGCTGAACCAAGAGTGTACCCCTTATGATCCGATGTGACACTT 1095
Db 121 CTCGTATGGGAGCATGCTGAACCAAGAGTGTACCCCTTATGATCCGATGTGACACTT 180
QY 1096 TTGTTAGTGTCCATGAGAAATATGCTCAAGCTGAGAGAGTATGAGAGCACATTTTC 1155
Db 181 TTGTTAGTGTCCATGAGAAATATGCTCAAGCTGAGAGAGTATGAGAGCACATTTTC 240
QY 1156 CAACTTTGGAATGTTCTAGTGTACAGATCAGATCAGATCCCAAAAATGAGCAGTGGATG 1215
Db 241 CAACTTTGGAATGTTCTAGTGTACAGATCAGATCAGATCCCAAAAATGAGCAGTGGATG 300
QY 1216 AAGAGAATCTTCATAGGATGACGAGAGTACGATCTGATGTGACGAGAGTATG 1275
Db 301 AAGAGAATCTTCATAGGATGACGAGAGTACGATCTGATGTGACGAGAGTATG 360
QY 1276 AGGACAATCTCCATAGCCCATTTCTGTCCAGGACGACACAGAGTGGAGAGAGACA 1335
Db 361 AGGACAATCTCCATAGCCCATTTCTGTCCAGGACGACACAGAGTGGAGAGAGACA 420
QY 1336 TTGTGACCATGCTGACCCGTGGAAGTCTTTGAGCATGAGAAAGCAAACTCTTAGGGG 1395
Db 421 TTGTGACCATGCTGACCCGTGGAAGTCTTTGAGCATGAGAAAGCAAACTCTTAGGGG 480
QY 1396 AGGTTGAGATGTTGAGACAGCATGATATCGGTGGGATGCGAGCTTGTGAAAT 1455
Db 481 AGGTTGAGATGTTGAGACAGCATGATATCGGTGGGATGCGAGCTTGTGAAAT 540
QY 1456 GGTCAAGAGAGAGTGAAGATGTTGAGAAAGAGAGTCTTCAAAAGAGTCTACTTGC 1515
Db 541 GGTCAAGAGAGAGTGAAGATGTTGAGAAAGAGAGTCTTCAAAAGAGTCTACTTGC 600
QY 1516 ACCAAGAGAGAGTCTTCAAGAGAGGAGTCTTCAACTTCCCGGTGGGGG 1575
Db 601 ACCAAGAGAGAGTCTTCAAGAGAGGAGTCTTCAACTTCCCGGTGGGGG 660
QY 1576 ATGTTTGTAGAGGTAGTGTGATGATGCTGCTCTTTAGTAAAGTCAAGTCACTTT 1635
Db 661 ATGTTTGTAGAGGTAGTGTGATGATGCTGCTCTTTAGTAAAGTCAAGTCACTTT 720
QY 1636 TCTCAAGAGGCTTGTGTAACCAAGCATGTGAGATGTCAGTGTGACCATCTGAGG 1695
Db 721 TCTC-AAGGCTTGTGTAACCAAGCATGTGAGATGTCAGTGTGACCATCTGAGG 778

QY 1696 TAGCTGCCAAGGTTCAAG-TTGGAAGATTTTGAACCTGAGTGAAGCGTCCCTG 1754
Db 779 TAGCTGCCAAGGTTCAAGTTTGGAAGATTTTGAACCTGAGTGAAGCGTCCCTG 838
QY 1755 TTAGTCG 1761
Db 839 TTAGTCG 845

RESULT 4
BZ989776/c
LOCUS BZ989776 795 bp DNA linear GSS 25-MAR-2003
DEFINITION PUGDY07TD ZM 0.6_1.0 KB Zea mays genomic clone ZM081A362B13,
genomic survey sequence.
ACCESSION BZ989776
VERSION BZ989776.1 GI:29229913
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 795)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished
Other_GSSs: PUGDY07TB
COMMENT Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white@lwtigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source location/Qualifiers
1..795
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM081A362B13"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

BASE COUNT 197 a 238 c 143 g 217 t

QY 817 TCTCAGTGAATGTTCCCTTCTTCTGAAAGGTTGAGGTTGAGAGACACTTCATTG 876
Db 740 TCACAGTGAATGTTCCCTTCTTCTGAAAGGTTGAGGTTGAGAGACACTTCATTG 681
QY 877 AAGATACATCATTTGACCTGCCACCGAGCAGCCGATGATCTTTACTGACGGTGATA 936
Db 680 AAGATACATCATTTGACCTGCCACCGAGCAGCCGATGATCTTTACTGACGGTGATA 621
QY 937 AGAACAATCACACTTTATGGGCTGAAGAGCCAGTCAATGATTGCTCGACCTCTA 996
Db 620 AGAACAATCACACTTTATGGGCTGAAGAGCCAGTCAATGATTGCTCGACCTCTA 561
QY 997 AGGAGCCCATATGCTTGAAGTGTCTTCTTCTGATCTGCTCATGAGAGCATGTGA 1056
Db 560 AGGAGCCCATATGCTTGAAGTGTCTTCTTCTGATCTGCTCATGAGAGCATGTGA 501
QY 1057 ACCAGAGTACCCCTTATGATCCGATGTGACACTTTTGTAGTGTCCATGAGAATA 1116
Db 500 ACCAGAGTACCCCTTATGATCCGATGTGACACTTTTGTAGTGTCCATGAGAATA 441


```
FEATURES
Source
    location/Qualifiers
    1..817
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMB1362B13"
    /clone_lib="ZM_0.6.1.0 KB"
    /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
    Cot selected genomic DNA library"

BASE COUNT      232 a      145 c      203 g      237 t
ORIGIN

Query Match      19.9%; Score 553.2; DB 29; Length 817;
Best Local Similarity 99.5%; Pred. No. 2.4e-56;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      817 TCTCAGGTGAATTGTCCTTCTTCTCGAAGGGTTGAGGTTGAGAGACACTTCATTG 876
      |||
Db      260 TCACAGGTGAATTGTCCTTCTTCTCGAAGGGTTGAGGTTGAGAGACACTTCATTG 319

QY      877 AAGAGTACATCATTTGAGACCTGCCACCGAGGAGCCGATGATCTTGTACTGACGGTGATA 936
      |||
Db      320 AAGAGTACATCATTTGAGACCTGCCACCGAGGAGCCGATGATCTTGTACTGACGGTGATA 379

QY      937 AGGAACAATCACACTTTATGGGCTGAAGAAGGCCAGTCATGATTGCTGCACCTCTA 996
      |||
Db      380 AGGAACAATCACACTTTATGGGCTGAAGAAGGCCAGTCATGATTGCTGCACCTCTCA 439

QY      997 AGGAGCCCATCATGCTTGAAGTGTGCTTCTTCTTCATCTCGTCATGGAGCATGTGA 1056
      |||
Db      440 AGGAGCCCATCATGCTTGAAGTGTGCTTCTTCTTCATCTCGTCATGGAGCATGTGA 499

QY      1057 ACCAGAGTGTACCCCTTATGATGCCGATGTGACACTTTTGTAGTGTCCATGAGATA 1116
      |||
Db      500 ACCAGAGTGTACCCCTTATGATGCCGATGTGACACTTTTGTAGTGTCCATGAGATA 559

QY      1117 TGCTCAAGCTGAGGAAGTATGAGGAGCACATTGTTCCAACTTTGGAAGTATGTCA 1176
      |||
Db      560 TGCTCAAGCTGAGGAAGTATGAGGAGCACATTGTTCCAACTTTGGAAGTATGTCA 619

QY      1177 GTGTCAAGATCAGATGCCAAATAGACAGTGGATGAGAGATCTTCATAGGATG 1236
      |||
Db      620 GTGTCAAGATCAGATGCCAAATAGACAGTGGATGAGAGATCTTCATAGGATG 679

QY      1237 ACCAGAGTACGATCTGATGTGACGAGGTGACTATGAGACAATCTCCATAGCCCAT 1296
      |||
Db      680 ACCAGAGTACGATCTGATGTGACGAGGTGACTATGAGACAATCTCCATAGCCCAT 739

QY      1297 TGCTGTCCAGGACGACAGGTGCGGAAGGGAAGGACATTTGCAACCATGTCACCGTG 1356
      |||
Db      740 TGCTGTCCAGGACGACAGGTGCGGAAGGGAAGGACATTTGCAACCATGTCACCGTG 799

QY      1357 GAAGTGTCTTGAGCATGA 1374
      |||
Db      800 GAAGTGTCTTGAGCATGA 817

RESULT 7
LOCUS      BQ752737      829 bp      mRNA      linear      EST 23-JUL-2002
DEFINITION WHE4118_E11_I22ZS wheat salt-stressed root cDNA library Triticum
aestivum cDNA clone WHE4118_E11_I22, mRNA sequence.
ACCESSION  BQ752737
VERSION    BQ752737.1  GI:21930519
KEYWORDS   EST
SOURCE     Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticaceae; Triticum.
REFERENCE  1 (bases 1 to 829)
AUTHORS   Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K., Dvorak,J.
```

```
TITLE
JOURNAL
COMMENT
    'Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
    The structure and function of the expressed portion of the wheat
    genomes - Salt-stressed root cDNA library
    Unpublished
    Contact: Olin Anderson
    US Department of Agriculture, Agriculture Research Service, Pacific
    West Area, Western Regional Research Center
    800 Buchanan Street, Albany, CA 94710, USA
    Tel: 5105595773
    Fax: 5105595818
    Email: oandersn@pw.usda.gov
    Sequences have been trimmed to remove vector sequence and low
    quality sequence with phred score less than 20
    Seq primer: SK primer.

FEATURES
Source
    location/Qualifiers
    1..829
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="Chinese Spring"
    /db_xref="taxon:4565"
    /clone="WHE4118_E11_I22"
    /issue_type="Roots"
    /dev_stage="Full tillering"
    /lab_host="E. coli SOLR"
    /clone_lib="wheat salt-stressed root cDNA library"
    /note="Vector: Lambda Uni-ZAP XR, excised phagemid
    Bluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Hydroponic
    plants grown to full tillering stage were treated with 150
    mM NaCl for either 12 hours or 7 days. Root tissues of the
    plants subjected to both types of treatment were collected
    separately at University of California, Davis (E. Akhunov
    and K. Deal in J. Dvorak's lab). Total RNA was prepared
    separately from the two samples (12h and 7day treatments),
    and equal amount of RNA was then pooled. PolyA RNA was
    purified from the pooled RNA, a cDNA library was made, and
    the cDNA clones were in vivo excised to give pBluescript
    SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)
    at the University of California, Davis. Colony plating,
    plasmid DNA preparations and DNA sequencing were performed
    in the OD Anderson lab (all other authors)."
```

```
BASE COUNT      177 a      199 c      215 g      238 t
ORIGIN

Query Match      19.1%; Score 531; DB 13; Length 829;
Best Local Similarity 80.6%; Pred. No. 1e-53;
Matches 647; Conservative 0; Mismatches 150; Indels 6; Gaps 2;

QY      1466 GGAAGGTGAGATGTTAGAAGAAGAGGTGTTCAAAAGAGTCTACTTGCAACAAGAGG 1525
      |||
Db      27  GAAAGGCGAGATGGCAAGAAGAGAGGCTTCAAAAGATCTACTTGCAACAAGAGG 86

QY      1526 AGTTCCTGGCTCAAGAGGGGCTCAATTGTTCACTTCCCGGTGTCGCGCATG--TTT 1582
      |||
Db      87  GGTGGCCGACTCAAGAGGGGCTCTGTTGTTCACTTCTGTCGGGGTATGCCACCA 146

QY      1583 TGAGGGTAGTGAATTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGACTTTTCTCAA 1642
      |||
Db      147 AGGGGGCAGTGGGTTTATACACGCTGCTGCTTGTGTAAGCCACTGGGCTTTTACTCCAA 206

QY      1643 GGGCTCTGTGAACGACGAT--GTCAGATGCTGCCATGTTCAACCATCTGAGGTAGC 1699
      |||
Db      207 GGATCTTATGAAGAGCGTATGCGGCCGTCAGCCATGATCCATCAATGGAAGCAGC 266

QY      1700 TGCCAAAGTTACAGTTGGAAGATTTGTTGAACCTGAGTGAGGCGTCCTGTAGT 1759
      |||
Db      267 TCCCAAGGTTCAATCTGGAAGAAGATCTGTTGAACCTGCTGAGGCGTGAATGTTGCT 326

QY      1760 CGGTGTGAATTCAGATCCTTCAACAGTTTGTGGAATAAACGGTGTCTTACTATAC 1819
      |||
Db      327 CGGGCTTGAATTCAGATGCTTCAAGCAGTTTGTGGAATAAATGAGTTCTTACTATAC 386

QY      1820 CCCACAAATCTTGAGCAAGCTGTGTGGCAGTTATCTTCCAAATTTGGTCTCAGCTC 1879
      |||
```


Db 387 TCCTCAATCTTGAGCAAGCTGTGGCGGTTCTTTTCCAACTTGCGCTCAGTTC 446
QY 1880 GGCATCAGCATTCATCTTGATCAGTCTCTCACTACTTAATGCTTCCCTTGCAATGG 1939
Db 447 AGCATCAGCGTCCATCTTGATCAGTCTCTCACCACCTTAATCATGCTCCCAAGCATGG 506
QY 1940 CTTGCCATGCTGCTTAAGATCTTCCGGAAGAGTTTGTCTGCTAGGACAAATTC 1999
Db 507 TGTAGCCATGAGACTTATGATATATCTGGAAGAAGTTTCTTACTGAGCACAATGCC 566
QY 2000 AATCTGATAGCATCTCTAGTATCTGCTGTGTGTCCAACTAATGATTGGGTACACT 2059
Db 567 CATCTGATAGCATCCCTAATGTTTGTGTGTGTCCAACTAATGATTGAGTACGGT 626
QY 2060 AGCCCATGCTTGTCTCTCCACCGCTCAGTGTATCTGCTCACTTCTGCTCTGCTATGG 2119
Db 627 GCCCCACGCTGTCTCTCCACAGTTAGCGTCAATGCTTCTGCTGCTTGTCTATGGG 686
QY 2120 ATTGGTCCCATCCCAACATTTTATGTGACAGATCTTCCAAACAGGTTGCTGGCT 2179
Db 687 CTTGGCCCAATCCCAACATCTATGTGACAGATTTTCCCAACAGAGTCCGTTGTGT 746
QY 2180 CTGTATGCTCATTTGTGCTTCAATCTGTGATCGAGATATCATGCTCACCCTACAGCT 2239
Db 747 CTGCATCGCTATTTGCGCCCTCACAATCTGTGATATGTGACATTAATTGTACTACAGCT 806
QY 2240 TCCTGTGATGCTGAATGCTATTG 2262
Db 807 GCCTGTGATGCTGAATGCTATTG 829

RESULT 8
BI992011/c 545 bp mRNA linear EST 24-OCT-2001
LOCUS 1020055D09.x4 1020 - Unigene II from Maize Genome Project Zea mays
DEFINITION cDNA, mRNA sequence.
ACCESSION BI992011
VERSION BI992011.1 GI:16377584
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 545)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plates: 1020055 row: D column: 09.
FEATURES
source
1..545
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="1020 - Unigene II from Maize Genome Project"
/note="This library represents the unique genes found in
the second round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
libraries 947 and 949. Contigs were assembled using
zmdBAssembler and 2 representatives from each contig were
selected for the Unigene set. All singlets were also
selected."

BASE COUNT 182 a 120 c 101 g 142 t
ORIGIN

Query Match 18.9%; Score 525; DB 12; Length 545;
Best Local Similarity 98.9%; Pred. No. 6.5e-53;
Matches 539; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 2178 CTCTGATTTGCCATTTGTGCTTACATTTGATCGGAGATATCATCGTCAACCTACAGC 2237
Db 545 CTCTGATTTGCCATTTGTGCTTACATTTGATCGGAGATATCATCGTCAACCTACAGC 486
QY 2238 CTTCTGTGATGCTGAATGCTATTGAGCTGGCGGTTTTCAGCATATATGACGTGTA 2297
Db 485 CTTCTGTGATGCTGAATGCTATTGAGCTGGCGGTTTTCAGCATATATGACGTGTA 426
QY 2298 TGCTTGATTTCCCTTTGTGTGCTCTTCCCTTAAGCTCCCTGAGACAAAGGGATGCCCTT 2357
Db 425 TGCTTGATTTCCCTTTGTGTGCTCTTCCCTTAAGCTCCCTGAGACAAAGGGATGCCCTT 366
QY 2358 GAGGTATTACCGAATCTTTGACAGTTGGTGGGAAGCAAGCGGCTGCAAAAGCCTAA-TT 2416
Db 365 GAGGTATTACCGAATCTTTGACAGTTGGTGGGAAGCAAGCGGCTGCAAAAGCCTAA-TT 306
QY 2417 TCTTTGATACCTTTGTGTGCAACTATTTGACACTGTAAGTATAGAACTTGAAGGGTTTAC 2476
Db 305 TCTTTGATACCTTTGTGTGCAACTATTTGACACTGTAAGTATAGAACTTGAAGGGTTTAC 246
QY 2477 CAAGAAGCTCGGAATTAATCTTTGATTTGTGTAATGTTAAGGGAACGAACATCTGCTC 2536
Db 245 CAAGAAGCTCGGAATTAATCTTTGATTTGTGTAATGTTAAGGGAACGAACATCTGCTC 186
QY 2537 ATGCTCTCAACCGTAATAAAGAGTCCCTCAATGCGCAATAGAGTCTGTAAGTTGTCA 2596
Db 185 ATGCTCTCAACCGTAATAAAGAGTCCCTCAATGCGCAATAGAGTCTGTAAGTTGTCA 126
QY 2597 ATGTCATTTACCATATGTTTACCTATTGTAAGTATTAAGTCAAGCTATTCACGC 2656
Db 125 ATGTCATTTACCATATGTTTACCTATTGTAAGTATTAAGTCAAGCTATTCACGC 66
QY 2657 TGGTTGTTGCTAGAAATCTTTAGAACAAAGATGAATGATCTGATCTGATGTTATAATA 2716
Db 65 TGGTTGTTGCTAGAAATCTTTAGAACAAAGATGAATGATCTGATCTGATGTTATAATA 6
QY 2717 TTCAA 2721
Db 5 TTCAA 1

RESULT 9
BM325261 572 bp mRNA linear EST 04-JAN-2002
LOCUS PIC1_42_E11.b1 A002 Pathogen-infected compatible 1 (PIC1) Sorghum
DEFINITION bicolor cDNA, mRNA sequence.
ACCESSION BM325261
VERSION BM325261.1 GI:18064398
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 572)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
Sudman,M. and Pratt,L.H.
TITLE An EST database from Sorghum: plants infected with a compatible
pathogen
JOURNAL Unpublished
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector, and regions
below Phred quality 16. The threshold for highest quality sequence

Db	Accession	Sequence	Length
QY	2149	CAGAGATCTTTCCAACCAAGGCTTCTGTCCTCTGTATTCGCAATTTGTCCTTACATCT	2208
Db	121	CAGAGATCTTTCCAACCAAGGCTTCTGTCCTCTGTATTCGCAATTTGTCCTTACATCT	180
QY	2209	GGATCGGAGATATCATCGTCACTTACAGCCTTCTGTGATGCTGAATGCTATTTGACTGG	2268
Db	181	GGATCGGAGATATCATCGTCACTTACAGCCTTCTGTGATGCTGAATGCTATTTGACTAG	240
QY	2269	CGGCTGTTTCAGCATATATGCAAGTGTATGCTTGAATTTCTTGTGCTTCTCTTA	2328
Db	241	CAGGTGTTTTCAGCATATATGCAAGTGTATGCTTGAATTTCTTGTGCTTCTCTTA	300
QY	2329	AGGTCCTTGAGACAAAGGGATGCCCCCTTGAGGTTATTAACCGAATCTTTGACGTTGGTG	2388
Db	301	AGGTCCTTGAGACAAAGGGATGCCCCCTTGAGGTTATTAACCGAATCTTTGACGTTGGTG	360
QY	2389	CGAAGCAAGCGGCTGCAAAAGCCTAATTTCTTGTGACCTTGTGTGCAAC-TATTGAC	2447
Db	361	CGAAGCAAGCGGCTGCAAAAGCCTAATTTCTTGTGACCTTGTGTGCAAC-TATTGAC	420
QY	2448	TGTAAGTTAGAACTTGAAGGGTTTACCAAGAGCTCGGAGAAATTAATCTTGGATTGT	2507
Db	421	TGTAAGTTAGAACTTGAAGGGTTTACCAAGAGCTCGGAGAAATTAATCTTGGATTGT	477
QY	2508	GTAATGTTAAGGAGACGACATCTGCTCATGCTCTCTCAAACGGTAAAGAGAGCTCCTC	2567
Db	478	GTAATGTTAAGGAGACGACATCTGCTCATGCTCTCTCAAACGGTAAAGAGAGCTCCTC	536
QY	2568	AATGCAATAGAGAGCTGTTAAGTGTCAATGCTATTAATTAATGCTTACCATATGTTT	2626
Db	537	AATGCAATAGAGAGCTGTTAAGTGTCAATGCTATTAATTAATGCTTACCATATGTTT	596
QY	2627	TACTGTATTATAGTCAAGCTATTTCAACGCTGTTGTTGCTAGAATCTTTAGACAAG	2686
Db	597	TACTGTATTATAGTCAAGCTATTTCAACGCTGTTGTTGCTAGAATCTTTAGACAAG	652
QY	2687	ATGATAATGATCTGATCTGATGTTAATATTAATTTCAAAATCTCAAAATAA	2733
Db	653	ATGATAATGATCTGATCTGATGTTAATATTAATTTCAAAATCTCAAAATAA	697

RESULT 11
BI074258
LOCUS
DEFINITION
IP1_13_H11.g1_A002 Immature panicle 1 (IP1) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION
BI074258
VERSION
BI074258.1 GI:14512915
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 664)
AUTHORS
Klein,R.R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt
L.H.
TITLE
An EST database from Sorghum: developing preanthesis panicles
JOURNAL
Unpublished
COMMENT
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with PolyMix or T7
sequencing primer, are presented as the reverse complement.
Seq primer: T7
High quality sequence start: 112

High quality sequence stop: 664
POLYA=Yes.

FEATURES
Source
Location/Qualifiers
1. .664
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultiivar="BTx623"
/db_xref="taxon:4558"
/clone_lib="Immature panicle 1 (IP1)"
/note="Organ: Developing preanthesis panicles; Vector: Bluescript II SK(-) from lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

BASE COUNT
165 a 141 c 144 g 214 t
ORIGIN

Query Match 16.9%; Score 469.6; DB 12; Length 664;
Best Local Similarity 87.2%; Pred. No. 2e-46;
Matches 586; Conservative 0; Mismatches 74; Indels 12; Gaps 6;

QY	2077	CCACCGTCAGTGTATTCGCTCTACTCTGCTGCTTGTATGGATTGGTCCCAATCCCA	2136
Db	1	CCACAGTCAGTGTATCATCTACTCTGCTGCTTGTATGGATTGGTCCCAATCCCA	60
QY	2137	ACATTTATGTGCAGAGATCTTCCACACGAGGTTGGCCCTGTATTTGCCATTTGTG	2196
Db	61	ACATTTATGTGCAGAGATCTTCCACACGAGGTTGGCCCTGTATTTGCCATTTGTG	120
QY	2197	CCTTACATTTGATCGGAGATATCATCTGCTACCTACAGCCCTTCTGTATGCTGAATG	2256
Db	121	CCTGACGTTTGGATTGGAGACATCATGTGTCACCTACAGCCCTTCTGTATGCTGAATG	180
QY	2257	CTATTGGACTGGCGGTTTTCAGCATATATGCACTGCTATGCTGATTTCTTGTGT	2316
Db	181	CTATTGGACTGAGGTTTTCAGCATATATGCACTGCTATGCTGATTTCTTGTGT	240
QY	2317	TGCTCTTCCCTTAAGTCCCTGAGACAAAGGGGATGCCCTTGAGGTATTACCGAATCT	2376
Db	241	TGCTCTTCCCTTAAGTCCCTGAGACAAAGGGGATGCCCTTGAGGTATTACCGAATCT	300
QY	2377	TTGCAAGTGGTGGCAAGCAAGCGGCTGCAAAAGCCTAATTCTTGGTACCTTGTGTGC	2436
Db	301	TTGCAAGTGGTGGCAAGCAAGCGGCTGCAAAAGCCTAATTCTTGGTACCTTGTGTGC	360
QY	2437	AAC-TATTGACATGTAAGTTAGAACTTGAAGGGGTTCAACCAAGAAGCTCGAGAAATTA	2495
Db	361	AACTTTTGGACATGTAAGTTAGAACTTGAAGGGGTTCAACCAAGAAGCTCTGGGA--A	417
QY	2496	CTTTGATTTGTGTAATGTTAAGGGAACGAACATCTGCTCATGCTCTCAACGGTAA	2555
Db	418	CTTTGATTTGTGTAATGTTAAGGGAACGAACATCTGCTCATGCTCTCAACGGTAA	476
QY	2556	AAAGAGTCCCTCAATGGCAAAATAGGAGTCTTAAGTGTCAATGTCATTTACCAATATGTT	2615
Db	477	CAAGACTCCCTCGATGGCAAAAGAGCAGTGTTAAGTGTCAATGTTATTTACCAATATGTT	536
QY	2616	-TTACCTATTTGTAATGTTATTAAGTCAAGCTAATTCACGCTGTTGTTGCTAGAAATC	2674
Db	537	GTTACCTATTTGTAATGTTATTAAGTCAAGCTAATTCACGCTGTTGTTGCTAGAAATC	593
QY	2675	TTTGAACAAGAATGATATGATCTGATCTGATGTTAATATTTCAAAATCTCAAAATAAG	2734
Db	594	TTTGAACAAGAATGATATGATCTGATCTGATGTTAATATTTCAAAATCTCAAAATAAG	650
QY	2735	AAAATATCGTTT 2746	
Db	651	AAAATATCGTTT 662	

RESULT 12
BG356047
LOCUS

653 bp mRNA linear EST 17-DEC-2001

DEFINITION EML_19_F09.g1_A002 Embryo 1 (EML) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BG356047 GI:15724666
VERSION BG356047.2
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 653)
AUTHORS Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
TITLE An EST database from Sorghum: developing embryos
JOURNAL Unpublished
COMMENT On Mar 6, 2001 this sequence version replaced gi:13238033.
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20. Three-prime sequences, which are obtained with PolyTmix or
T7 sequencing primer, are presented as the reverse complement.
Seq primer: PolyTmix
High quality sequence start: 38
High quality sequence stop: 612
POLYA=No.
FEATURES
source 1.653
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Embryo 1 (EML)"
/note="Organ: Embryos germinated for 24 hr; Vector:
pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:
EcoRI; The library was made from poly-A RNA in the cloning
vector lambda Zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT 153 a 142 c 149 g 209 t
ORIGIN
Query Match 16.8%; Score 466.2; DB 10; Length 653;
Best Local Similarity 86.8%; Pred. No. 5e-46;
Matches 572; Conservative 0; Mismatches 78; Indels 9; Gaps 5;
QY 2032 TGTCCAATCTAATTGATTGGGTACACTAGCCCATCTTTGCTCTCCACCGTCAGTGTTA 2091
DB 1 TGTCCAATGTGATTGACTGGGTACAGTGGCCCATCTGCGCTCTCCACAGTCAGTGTCA 60
QY 2092 TCGTCTACTTCTGCTGCTTGTATGGGATTGGTCCATCCCAACATTTTATGTGCAG 2151
DB 61 TCATCTACTTCTGCTGCTTGTATGGGATTGGTCCATCCCAACATTCTATGTGCAG 120
QY 2152 AGATCTTTCCAACCAAGGTTCTGGGCTCTGATTTGCCATTGTGCTTTACATTCTGGA 2211
DB 121 AGATCTTTCCAACCAAGGTTCTGGGCTCTGATTTGCCATTGTGCTTTGACGTTTGA 180
QY 2212 TCGGAGATATCATCTGTCACTACAGCCTTCTGATGCTGTAATGCTATTTGACTGGCGG 2271
DB 181 TTGAGACATCATTTGTCACTACAGCCTTCTGATGCTGTAATGCTATTTGACTAGCAG 240
QY 2272 GTGTTTTCAGCATATATGAGTGTATGCTTGTGTTGTTGCTCTTCTTTAAGG 2331
DB 241 GTGTTTTCAGCATATATGAGTGTATGCTTGTGTTGTTGCTCTTCTTTAAGG 300
QY 2332 TCCCTGAGACAAAGGGGATGCCCTTGAGTTATTAACGAATCTTTGACAGTTGGTCCGA 2391
DB 301 TTCCTGAGACAAAGGGGATGCCCTTGAGTTATTAACGAATCTTTGACAGTTGGTCCGA 360
QY 2392 AGCAAGCGGTGCAAAAGCCTAATTCTTTGTGTACTTTGTGTGCAAC-TATTGCACTGT 2450

|||||
DB 361 AGCAAGCGGTGCAAAAGCCTAATTCTTTGTGTACTTTGTGTGCACTGT 420
QY 2451 AAGTTAGAACTTGAAGGGGTTTACCAAGAAGCTCGAGAATTACTTTGGATTGTGTA 2510
DB 421 GAATTATAACTTGAAGGGATCTCACCAGAAGCTCTGGA--ACTTTGATTTGTATA 477
QY 2511 AATGTTAAGGAACGAACATCTGCTCATGCTCTCAACCGGTAAAGAGTCCCTCAAT 2570
DB 478 AATGTGAGGGAACAGATATCTGCTCATGCTCTCGCAAGC-GACAGCAAGACTCCCTCAT 536
QY 2571 GGCAATAGAGTCTTAAAGTTGTCATGCTATTACCATATGTT-TTACCTATTGTGTA 2629
DB 537 GGCAAGAGAGTGTAAAGTTGTCATGCTATTACCATATGTTGTTACCTATTAGTAC 596
QY 2630 TGTATTAACTCAAGCTATTCAACGCTGCTTGTGCTAGAATCTTTAGAACAAAGAT 2688
DB 597 TGTATT--AGTCAAGATATTACGACACTGTTGTTAGTAGAGATGTTAGAACAAAGAT 652
RESULT 13
BM329724 661 bp mRNA linear EST 04-JAN-2002
LOCUS PIC1_42_E11.g1_A002 Pathogen-infected compatible 1 (PIC1) Sorghum
DEFINITION bicolor cDNA, mRNA sequence.
ACCESSION BM329724
VERSION BM329724.1 GI:18068861
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 661)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
Sudman,M. and Pratt,L.H.
TITLE An EST database from Sorghum: plants infected with a compatible
pathogen
JOURNAL Unpublished
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector, and regions
below Phred quality 16. The threshold for highest quality sequence
is 20. Three-prime sequences, which are obtained with PolyTmix or
T7 sequencing primer, are presented as the reverse complement.
Seq primer: T7
High quality sequence start: 57
High quality sequence stop: 661
POLYA=Yes.
FEATURES
source 1.661
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with
Colletotrichum graminicola"
/clone_lib="Pathogen-infected compatible 1 (PIC1)"
/note="Vector: pBluescript II SK(-) from Lambda Zap II;
Site_1: XhoI; Site_2: EcoRI; Four-week-old sorghum
seedlings were sprayed with spore suspension prepared from
3-week-old FRM42I, a sorghum isolate of the anthracnose
pathogen Colletotrichum graminicola. Inoculated plants
were kept in a 25 C dark growth chamber with 100% relative
humidity for 24 hr, followed by 12/12 hr of light/dark
cycle at 25 C with 90% relative humidity for another 24
hr. All leaves were harvested and quick frozen with liquid

SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticaceae; Triticum.
1 (bases 1 to 703)
AUTHORS Anderson,O.D., Chao,S., Crossman,C., Langridge,P., Lazo,G.R., Pham
,J., Rausch,C.J., Sutton,T., Woo,J. and Wilson,C.
TITLE The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
JOURNAL Unpublished
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: T7 primer.
FEATURES
source
1..703
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4048_B04_D08"
/tissue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="Wheat meiotic anther cDNA library"
/note="Vector: pSPORT1, Site_1: SalI; Site_2: NotI; Plants
were grown in a glasshouse. Anther meiotic stage was
determined by removing anthers from individual primary
florets. One anther was sacrificed for microscopic staging
, and if determined to be between (and including) meiotic
stages pre-meiosis and metaphase I, the remaining two
anthers were collected and pooled for library
construction. The tissue, total RNA, and poly(A) RNA were
prepared, cDNA synthesised, and directionally ligated into
pSPORT1 by Tim Sutton in the P Langridge lab at the
Department of Plant Science, University of Adelaide, Waite
Campus, Australia. Average insert size 1.5Kb. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 139 a 183 c 165 g 216 t
ORIGIN
Query Match 16.6%; Score 460; DB 14; Length 703;
Best Local Similarity 81.6%; Pred. No. 2.6e-45;
Matches 545; Conservative 0; Mismatches 120; Indels 3; Gaps 1;
QY 1602 CATGCTGCTGCTTTAGTACGTCAGCACTTTCTCAAAGGGTCTTGTAACCAAGC 1661
DB 36 CACGCAGCTGCTTGTGTAAGCCACTCGGCTCTTACTCAAGATCTTATGAAGAGCGT 95
QY 1662 ATGTCAGATGCT--GCCATGGTCAACCATCTGAGGTAGCTGCCAAAGTTCAAGTTGG 1718
DB 96 ATGGCGGCTGTCAGCCATGATTCATCCATTGGAGCGAGCTCCCAAGTTCAATCTGG 155
QY 1719 AAAGATTGTTGTAACCTGAGTGAAGCGCGTCCCTGTTAGTCGGTGTGGAATTCAGATC 1778
DB 156 AAAGATCTGTTGAACCTGAGTGAAGCGCGTGCATTTGTCGTCGCGCTTGAATTCAGATG 215
QY 1779 CTTCAACAGTTTGTGTAATTAAGCGTGTCTGTACTATATACCCACAAATCTTGAGCAA 1838
DB 216 CTTCAAGCAGTTTGTGTAATTAAGAGTTCTCTACTATATCTCTCAAAATCTGAGCAA 275
QY 1839 GCTGCTGTGGCAGTTATCTTTCCAAATTTGGTCTCAGCTGGCATCAGCATCCATCTTG 1898
DB 276 GCTGCTGTGGCGGTTCTTCTTTCCAAATCTTGCCCTCAGTTCAAGCATCAGCATCCATCTTG 335

QY 1899 ATCAGTCTCTCACTACCTTACTATATGCTTCCTTGATTTGGCTTTGCCATGCTGCTTATG 1958
DB 336 ATCAGTCTCTCACCACTTACTATCATGCTCCCAAGCATTTGGTAGCCATGAGACTTATG 395
QY 1959 GATCTTTCCGGAAGAGTTTGTGCTGCTAGGCACAAATTCATCTTGATATGATCTCTTA 2018
DB 396 GATATATCTGGAAGAGGTTTCTGCTACTAGGCACAAATTCATCTTGATATGATCTCCCTA 455
QY 2019 GTTATCTGTTGTGTCCTCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCC 2078
DB 456 ATGTTTGTGTTGTGTCCTCAATGTTATCAACTTGAATGAGTGCCTCCACGCTGTCTCTCC 515
QY 2079 ACCGTCACTGTTATCGTCTACTTCTGCTGCTTGTATGGGATTTGTCCTCCATCCCAAC 2138
DB 516 ACAGTTAGCGTCACTGCTACTTCTGCTGCTTGTATGAGGCTTTGGCCGATCCCAAC 575
QY 2139 ATTTATGTGACAGATCTTTCCACCAAGGTTGCGCTCTGTATGCGATTTGTGCC 2198
DB 576 ATTCTATGTGACAGATTTTCCCAAGGCGGTGTCTGCATCGCTATTTGCCGCC 635
QY 2199 TTTACATCTGATCGAGATATCATCGTCACTACACGCTTCTGTGATGCTGAATGCT 2258
DB 636 CTCACATCTGATTTGTGACATATTTGTTACTTACACGCTGCTGTGATGCTGAATGCT 695
QY 2259 ATTGACT 2266
DB 696 ATTGCTCT 703

Search completed: January 5, 2004, 15:47:47
Job time : 5668 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 09:46:20 ; Search time 10059 Seconds
(without alignments)
11293.969 Million cell updates/sec

Title: US-10-051-909-31

Perfect score: 2777
Sequence: 1 caccgggttagattcggagc.....aaaaaaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2200.2	79.2	2824	6	AR208565	AR208565 Sequence
2	1642	59.1	2665	8	AY165599	AY165599 Saccharum
3	1075.4	38.7	2614	8	HVU534445	AJ534445 Hordeum v
4	946.8	34.1	137140	2	AP005756	AP005756 Oryza sat
5	940.2	33.9	2601	6	AR208568	AR208568 Sequence
6	932.4	33.6	2570	8	AY094465	AY094465 Arabidops
7	875.8	31.5	2190	6	AX506620	AX506620 Sequence
8	875.8	31.5	2190	8	ATH532570	AJ532570 Arabidops
9	806.2	29.0	2516	8	HVU534446	AJ534446 Hordeum v
10	700	25.2	2205	6	AX412656	AX412656 Sequence
11	700	25.2	2205	6	AX507559	AX507559 Sequence
12	691.2	24.9	1487	6	AR208571	AR208571 Sequence
13	672.2	24.2	2426	8	ATSUGTRPR	Z50752 A.thaliana
14	646.2	23.3	2190	8	ATH532571	AJ532571 Arabidops
15	637.8	23.0	135406	8	AC136843	AC136843 Oryza sat
16	599.2	21.6	86710	8	ATF23E12	AL022604 Arabidops
17	599.2	21.6	197859	8	ATCHRIV83	AL161587 Arabidops
18	594.6	21.4	1692	6	AR208569	AR208569 Sequence
19	585.4	21.1	142114	8	AC073166	AC073166 Oryza sat
20	585.4	21.1	300957	8	AE017116	AE017116 Oryza sat
21	518.2	18.7	100693	2	AC121239	AC121239 Medicago
22	402.8	14.5	103192	8	AC007369	AC007369 Arabidops
23	402.8	14.5	143879	8	AC069251	AC069251 Genomic s
24	394.4	14.2	1009	6	AR208572	AR208572 Sequence
25	386.6	13.9	94349	8	ATF26013	AL133452 Arabidops
26	318.8	11.5	120538	2	AP004082	AP004082 Oryza sat
27	315.8	11.4	114918	8	AP004945	AP004945 Lotus jap
28	311.8	11.2	87592	8	AC099739	AC099739 Oryza sat
29	311.8	11.2	127098	8	AC144426	AC144426 Oryza sat
30	311.8	11.2	154128	8	AP000615	AP000615 Oryza sat
31	276.2	9.9	86162	2	AC126786	AC126786 Medicago
32	276.2	9.9	122107	2	AC144482	AC144482 Medicago
33	249	9.0	145365	2	AC127019	AC127019 Medicago
34	227	8.2	870	6	AR208567	AR208567 Sequence
35	225.2	8.1	70311	8	AF128457	AF128457 Oryza sat
36	220.4	7.9	77605	8	AF119222	AF119222 Oryza sat
37	218.8	7.9	142852	8	AF161269	AF161269 Oryza sat
38	194.4	7.0	108553	8	AF527809	AF527809 Sorghum b
39	185	6.7	106246	8	AF528565	AF528565 Zea mays
40	185	6.7	131457	2	AC144718	AC144718 Zea mays
41	185	6.7	346296	6	AF090447	AF090447 Zea mays
42	143.4	5.2	155	6	AR249238	AR249238 Sequence
43	123.2	4.4	510	6	AR208570	AR208570 Sequence
44	120.4	4.3	137568	8	AC093178	AC093178 Oryza sat
45	120.4	4.3	302228	8	AE017070	AE017070 Oryza sat

ALIGNMENTS

RESULT 1
AR208565
LOCUS AR208565 2824 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6383776.
ACCESSION AR208565
VERSION AR208565.1 GI:21509752
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2824)
AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 1 07-MAY-2002;
FEATURES Location/Qualifiers

Source	1..2824	/organism="unknown"
BASE COUNT	644 a	649 c 745 g 775 t 11 others
ORIGIN		
Query Match	79.2%;	Score 2200.2; LB 6; Length 2824;
Best Local Similarity	90.2%;	Pred. No. 0;
Matches 2372;	Conservative 0;	Mismatches 253; Indels 6; Gaps 2;

Db	Query Match	Similarity	Score	LB	Length
133	CTTCACAGGAGCCGGTGACCTCGGACGATATCTTGAGGACAAAGATGTCGGGGCTGTC	100.0%	2200.2	6	2824
194	CTTGCCCGCCGCTGCGGAGTGGGGCGTAGATTCCGGCGCCATGGGGGGCGCGTGA	90.2%	2200.2	6	2824
193	TTGTGCGCCATAGTGCCTCCATCGGCAATCTATTGACAGGGGTGGCAATGCCACCATCG	89.8%	2200.2	6	2824
254	TGTGTGCCCATCGCGCCCTTATCGGCAATCTGCTGACAGGGCTGGGACAAATGCGACAATTG	89.8%	2200.2	6	2824
253	CAGCTGCTGTTCTGTATATAAAGAGAAATTTCAATTGCAAAATGAGCCCACTGTGAGG	89.8%	2200.2	6	2824
314	CTGAGCCCGTCTCTGATACATAAAGAGAAATTTCAACCTGACAGCGAGCCTTGATCGAAG	89.8%	2200.2	6	2824
313	GACTAATTGTGTCATGTCACTTATCGCGCCACCATCGTTACTTACATTTCTCCGGCCAT	89.8%	2200.2	6	2824
374	GCCCTACGTCGCCCATGTTCCTCATTTGGGGCAACAGTCATCACACATCTCCGGGCCAA	89.8%	2200.2	6	2824
373	TATCAGACTCGATTGGCCGACGCCCTATGCTTATTCTCTCTTCAATTCTGTACTTCTCA	89.8%	2200.2	6	2824
434	GGGCTGACTCGGCTTGGTAGAGAGGCCCATGCTGTCGCCCTCGGCTGCTCTACTTCTGTC	89.8%	2200.2	6	2824
433	GCGGCTCATCATGCTATGCTCTCTTAATGCTATGCTCTGCTGTTGGACAGCTTCTGATG	89.8%	2200.2	6	2824
494	GTTGGCTGTGATGCTTTGGGCGCCAAATGTGTACATCTTGCTCTCTCGCAAGGCTCATTTG	89.8%	2200.2	6	2824
493	ATGATTTGGTATTGGCTTGGCTGTGACAGCTGTGCTTGTGATATTGAGAAATAGCCC	89.8%	2200.2	6	2824
554	ATGGTTGGTATCGGTTTGGCGGTGACACTTGTCTCTCTTACATCTCGAAACTGACAC	89.8%	2200.2	6	2824
553	C---TTCCGAGATTAGAGGTTTGTGAATACACTACCAATTCAGTGG--ATCAGAG	89.8%	2200.2	6	2824
614	CGCACAGANATTCTTGGGCTGTTNGAACACGTTGCCGAGTTCAATTGGGGTCAAGNGAG	89.8%	2200.2	6	2824
607	GAAATGTTCTGTATCATCTGATGTTGGGATGTCCTGTCGCCATCACCCGATTGGA	89.8%	2200.2	6	2824
674	GGAATGTTCTCTCTACTCATGTGTTGGGATGTCCTCATGCCCAACCTGATTGGA	89.8%	2200.2	6	2824
667	GAAATTAATGCTTGGTGTGCTCGCGATACTTCAATGTTCTTCTTGTGTTGACAAATATTTT	89.8%	2200.2	6	2824
734	GGCTCATGCTTGGAGTTCTGTGATCCGCTCACTTATNTACTTTGACTGACTGTCTTCT	89.8%	2200.2	6	2824
727	ATCTTCTGAAATCTCCAGATGCTGCTTGAACAAAGTCCGATGGCAGAGGCAAAAAAG	89.8%	2200.2	6	2824
794	ACTTGCTGAATCACCAAGGTGGCTTGNAGCAAAAGAGATGGCGGAGGCGAAGAGAG	89.8%	2200.2	6	2824
787	TGTTGCAAAAAGTTACGGGGGAAAGACATGCTCAGGTAATTTGCTTCTTCTCGAAG	89.8%	2200.2	6	2824
854	TGNTGCAAAAGGCTGCGGGAGAGAGATGTCTCANGGAGANGGCTTTCTAGTTGAAG	89.8%	2200.2	6	2824
847	GGTTGAGAGTTGAGAGACACTTCCATTGAAGATACATCATTTGACCTGACCAGAG	89.8%	2200.2	6	2824
914	GTTTGGGGGTCGGTAAGATACAGTATTTNAGATACATCATTTGACCTGCCACCGAGG	89.8%	2200.2	6	2824
907	CAGCGGATGATCTTGTACTGACGGTGAATAAGAAACAATCACACTTATGGGCTGAAG	89.8%	2200.2	6	2824
974	CAGCGGATGATCTTGTACTGACGGTGAATAAGAAACAATCACACTTATGGGCTGAAG	89.8%	2200.2	6	2824
967	AAGGCAATCATGATTTGCTCGACCTTCTTAAGGGAACCATCATGCTTGAAGTGTCTTT	89.8%	2200.2	6	2824
1034	AAGGCCAGTCATGATTTGCTCGACCTTCTTAAGGGAACCATCATGCTTGAAGTGTCTTT	89.8%	2200.2	6	2824
1027	CTTTTCATCTCTGTCATGGGAGCATGTTGAACCAAGTGTACCCCTTATGGATCCGATTG	89.8%	2200.2	6	2824
1094	CTTTTCATCTCTGTCATGGGAGCATGTTGAACCAAGTGTACCCCTTATGGATCCGATTG	89.8%	2200.2	6	2824

QY	1087	TGACAC	TTTTGGT	AGTGTCC	ATGAGAA	TATGCTT	CAAGCTG	GAGGA	AGTATG	AGGACA	1146				
Db	1154	TGACAC	TTTTGGT	AGTGTCC	ATGAGAA	TATGCTT	CAAGCTG	GAGGA	AGTATG	AGGACA	1213				
QY	1147	CATTG	TTTCCAA	CTTGG	AAGTATG	TTCA	GATCAG	ATGCCCC	AAATGAGC	1206					
Db	1214	CATTG	TTTCCAA	CTTGG	AAGTATG	TTCA	GATCAG	ATGCCCC	AAATGAGC	1273					
QY	1207	AGTGG	ATGAAGA	AACTT	CATAG	GGATG	ACGAG	GTACG	CATCTG	ATGTG	CAGAG	1266			
Db	1274	AGTGG	ATGAAGA	AACTT	CATAG	GGATG	ACGAG	GTACG	CATCTG	ATGTG	CAGAG	1333			
QY	1267	GTGAC	TATGAG	CAATCT	CATAG	CCCATG	CTGTCC	AGGCA	ACAGGTG	CGGAG	1326				
Db	1334	GTGAC	TATGAG	CAATCT	CATAG	CCCATG	CTGTCC	AGGCA	ACAGGTG	CGGAG	1393				
QY	1327	GGAAG	GACAT	TGTGC	ACCATG	GTCA	CCCGT	TGGA	GTGCTT	GACATG	AGAAG	1386			
Db	1394	GGAAG	GACAT	TGTGC	ACCATG	GTCA	CCCGT	TGGA	GTGCTT	GACATG	AGAAG	1453			
QY	1387	TCTTA	GGGAG	GGTGA	GATG	GTGTG	AGCAG	ACATG	ATTCG	GTGGG	GATGG	CAGCTT	1446		
Db	1454	TCTTA	GGGAG	GGTGA	GATG	GTGTG	AGCAG	ACATG	ATTCG	GTGGG	GATGG	CAGCTT	1513		
QY	1447	CTTGA	AAATG	GTCA	GAGAA	GGTGA	GAATG	GTAGAA	AGGAAG	GTGTT	CAAA	AGAG	1506		
Db	1514	CTTGA	AAATG	GTCA	GAGAA	GGTGA	GAATG	GTAGAA	AGGAAG	GTGTT	CAAA	AGAG	1573		
QY	1507	TCTAC	TTGC	ACCAAG	AGGAG	TTC	CTG	GCTCA	AGAGGG	CTCAT	TGTT	CTACT	CCCG	1566	
Db	1574	TCTAC	TTGC	ACCAAG	AGGAG	TTC	CTG	GCTCA	AGAGGG	CTCAT	TGTT	CTACT	CCCG	1633	
QY	1567	GTGTG	CGCAT	GTTTT	TGAG	GGGTAG	TGAG	TTTGT	ATCATG	CTGCTG	CTTTAG	TAAGT	CAGT	1626	
Db	1634	GTGTG	CGCAT	GTTTT	TGAG	GGGTAG	TGAG	TTTGT	ATCATG	CTGCTG	CTTTAG	TAAGT	CAGT	1693	
QY	1627	CAGCA	CTTT	CTCAA	AGG	STCTG	CTA	AA	CCG	CATG	TCAG	ATGCTG	CCATG	GTTC	1686
Db	1694	CAGCA	CTTT	CTCAA	AGG	STCTG	CTA	AA	CCG	CATG	TCAG	ATGCTG	CCATG	GTTC	1753
QY	1687	CATCT	GAGT	AGCTG	CCAA	AGG	STCTG	CTA	AA	GAAT	TTGTT	GAA	CTTGA	AGTG	1746
Db	1754	CATCT	GAGT	AGCTG	CCAA	AGG	STCTG	CTA	AA	GAAT	TTGTT	GAA	CTTGA	AGTG	1813
QY	1747	GTGCC	CTGT	AGTC	GGTGA	ATT	CAG	ATC	CTTCA	CAGTT	TGCTG	GAATA	ACG	GTG	1806
Db	1814	GTGCC	CTGT	AGTC	GGTGA	ATT	CAG	ATC	CTTCA	CAGTT	TGCTG	GAATA	ACG	GTG	1873
QY	1807	TTCTG	TACT	ATAC	CCCA	CAAA	TCTT	G	AGCA	AGCTG	GTG	CAGT	TAT	CTT	1866
Db	1874	TTCTG	TACT	ATAC	CCCA	CAAA	TCTT	G	AGCA	AGCTG	GTG	CAGT	TAT	CTT	1933
QY	1867	TTGGT	CTC	AGCTCG	GCAT	CAG	ATCC	ATCTT	GATC	AGTTCT	CTCA	CTAC	CTT	ACT	1926
Db	1934	TTGGT	CTC	AGCTCG	GCAT	CAG	ATCC	ATCTT	GATC	AGTTCT	CTCA	CTAC	CTT	ACT	1993
QY	1927	TTCC	TTG	CAT	TGG	CTT	TG	CCATG	CTG	CTTAT	GATCTT	CCG	AAG	AGT	1986
Db	1994	TTCC	TTG	CAT	TGG	CTT	TG	CCATG	CTG	CTTAT	GATCTT	CCG	AAG	AGT	2053
QY	1987	TAGG	CA	CAAT	CCA	ATCTT	GAT	AG	ATC	CTCTA	GTAT	TCTG	GTG	TCC	2046
Db	2054	TAGG	CA	CAAT	CCA	ATCTT	GAT	AG	ATC	CTCTA	GTAT	TCTG	GTG	TCC	2113
QY	2047	ATT	TGG	TAC	ACTAG	CCCATG	CTT	GCTC	TC	CA	CCGTC	AGTGT	TAT	CGT	2106
Db	2114	ATT	TGG	TAC	ACTAG	CCCATG	CTT	GCTC	TC	CA	CCGTC	AGTGT	TAT	CGT	2173
QY	2107	GCTT	CGT	TAT	GGG	ATTTG	GTCC	CA	TCC	CC	CA	CA	TTTAT	TG	2166
Db	2174	GCTT	CGT	TAT	GGG	ATTTG	GTCC	CA	TCC	CC	CA	CA	TTTAT	TG	2233
QY	2														

Db 2234 GGGTCGTGGCCTCTGTATGGCATTTGTGCCTTACATCTCGATCGAGATATCATCG 2293
QY 2227 TCACCTACAGCCTTCCTGTGATGCTGAATGCTATGGAAGTGGCGGTGTTTCAGCATAT 2286
Db 2294 TCACCTACAGCCTTCCTGTGATGCTGAATGCTATGGAAGTGGCGGTGTTTCAGCATAT 2353
QY 2287 ATGCACTGCTATGCTGTGATTTCTTTGTGCTGCTCTTAAAGTCCCTGAGACAAAGG 2346
Db 2354 ATGCAGTCGATGCTGTGATTTCTTTGTGCTGCTCTTAAAGTCCCTGAGACAAAGG 2413
QY 2347 GGATGCCCCCTTGAGGTATTAACGAATTTCTTGCAAGTTGGCGAAGCAAGCGGCTGCA 2406
Db 2414 GGATGCCCCCTTGAGGTATTAACGAATTTCTTGCAAGTTGGCGAAGCAAGCGGCTGCA 2473
QY 2407 AAGCCTAATTTCTTTGGTACCTTTGTGCAACTATGCACTGTAAGTTAGAACTTGAA 2466
Db 2474 AAGCCTAATTTCTTTGGTACCTTTGTGCAACTATGCACTGTAAGTTAGAACTTGAA 2533
QY 2467 GGGGTTTCACCAAGAGCTCGAGAAATTAATTTGATTTGTAATGTTAAGGAAAGCA 2526
Db 2534 GGGGTTTCACCAAGAGCTCGAGAAATTAATTTGATTTGTAATGTTAAGGAAAGCA 2593
QY 2527 ACATCTGCTCATGCTCCTCAACGGTAAAAAGAGTCCCTCAATGGCAATAGGAGTCTG 2586
Db 2594 ACATCTGCTCATGCTCCTCAACGGTAAAAAGAGTCCCTCAATGGCAATAGGAGTCTG 2653
QY 2587 TAAGTGTCAATGTCATTTACCATATGTTTACCTATTTGACTGTATTTAATGTAAGCAAG 2646
Db 2654 TAAGTGTCAATGTCATTTACCATATGTTTACCTATTTGACTGTATTTAATGTAAGCAAG 2713
QY 2647 TATTCACGCTGCTGTTGTGCTGAATCTTTAGAACAAAGATGATAATGATCTGATCTGA 2706
Db 2714 TATTCACGCTGCTGTTGTGCTGAATCTTTAGAACAAAGATGATAATGATCTGATCTGA 2773
QY 2707 TGTATATATATTCAAATCTCAATAAAGAAATATCGTTTCGAAAAAAA 2757
Db 2774 TGTATATATATTCAAATCTCAATAAAGAAATATCGTTTCGAAAAAAA 2824

RESULT 2
AY165599 2665 bp mRNA linear PLN 07-JUN-2003
LOCUS Saccharum hybrid cultivar putative sugar transporter type 2a mRNA,
DEFINITION complete cds.
ACCESSION AY165599
VERSION AY165599
KEYWORDS AY165599.1 GI:31505503
SOURCE Saccharum hybrid cultivar (sugarcane)
ORGANISM Saccharum hybrid cultivar

REFERENCE 1 (bases 1 to 2665)
AUTHORS Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.

TITLE Identification of a novel sugar transporter homologue strongly expressed in maturing stem vascular tissues of sugarcane by expressed sequence tag and microarray analysis

JOURNAL Plant Mol. Biol. 52, 371-386 (2003)
REFERENCE 2 (bases 1 to 2665)
AUTHORS Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.

TITLE Direct Submission
JOURNAL Submitted (17-OCT-2002) Division of Plant Industry, Commonwealth Scientific and Industrial Organisation, 120 Meiers Rd, Indooroopilly, QLD 4068, Australia

FEATURES
source 1..2665
/organism="Saccharum hybrid cultivar"
/mol_type="mRNA"
/db_xref="taxon:128810"
CDS 97..2334

/codon_start=1
/product="putative sugar transporter type 2a"
/protein_id="AA037640.1"
/db_xref="GI:31505504"
/translation="MSGALVAIVASIGNLLQGMNATISAAIYIKKEKLESEPTV
EGLIVAMSLIGATITITPSGPVSDWIGRLPMLILSSVLXFVSSLIMSPNVVYLLLA
RLINGFGVGLAVTLVPLIYSETAPPEIRGLNLTLPQFSSGGMFLSCMVFGLSLPS
PDWRIMLGLVALPSLFFPGULTIFYLPESPRLVSKGRMAEAKVLOLRGKDDVTGEM
ALLVEGLEVGDDTFIEBYIIGPAIEPADHDVADDERIILFPERGQSIARPSKGS
SMLGSVLSIGSRHGSVMNONPLTDPIVTLFGSVHENISLAGSMMRIVPSKFDMSIS
ITDHPKTEQWDEENLHRDNEEYATDGAASDYENNVSPLSRONTSAEGKIAHGH
RGSALGLRRRLSDEGEAATSTGIGGWQLAWKMSEREGEDKKEGSFKRIYLHOEG
VAGSRGVSLSLPGGDASEGKFIIHAALVSQALYPRDITEGRMAGPATMHPSEAA
AKVPSWRDLPEPGVRRALLVIGIQLQFAGINGVLVYTPQIMEQAGVAVLISNLGL
SSASASILISSVTALLMLPSIGLAMRLMDVSGRRFLLSTIPVLASLIVLVSNVIE
LGTVHAVLSTISVITYLCCPKMGFGPIPNILCAEFPTRRVRCITAIKIFWVDI
IVTYSLPVMLNAIGLEGVFGIYAVACALAFVYVYLKVPETKMPLEVITEFFAVGAKO
AVAKA"

misc_feature 97..156
/note="putative; transmembrane-region site"
misc_feature 115..774
/note="pfam00083; Region: sugar transporter"
misc_feature 238..297
/note="putative; transmembrane-region site"
misc_feature 313..372
/note="putative; transmembrane-region site"
misc_feature 406..462
/note="putative; transmembrane-region site"
misc_feature 511..570
/note="putative; transmembrane-region site"
misc_feature 589..654
/note="putative; transmembrane-region site"
misc_feature 658..1794
/note="Region: large cytoplasmic loop"
misc_feature 1624..2298
/note="pfam00083; Region: sugar transporter"
misc_feature 1672..1731
/note="putative; transmembrane-region site"
misc_feature 1813..1869
/note="putative; transmembrane-region site"
misc_feature 1897..1953
/note="putative; transmembrane-region site"
misc_feature 1969..2022
/note="putative; transmembrane-region site"
misc_feature 2089..2151
/note="putative; transmembrane-region site"
misc_feature 2191..2247
/note="putative; transmembrane-region site"

BASE COUNT 650 a 561 c 691 g 763 t
ORIGIN

Query Match 59.1%; Score 1642; DB 8; length 2665;
Best Local Similarity 79.4%; Pred. No. 0;
Matches 2003; Conservative 0; Mismatches 495; Indels 24; Gaps 4;

QY 167 GGAGACAAGATGTCGGGGCTGTTCTGTCCGCATAGTCGCCCTCCATCGGCAATCTATT 226
Db 87 GGGGAAAAACATGTCGGAGCTGCTCTTGTGCAATAGTACTTCGATCGGCAATCTGCT 146
QY 227 GCAGGGGTGGACAATGCCACCATCGCAGCTGCTGTTCTGTATATAAGAAAGAAATTCA 286
Db 147 TCAGGGGTGGACAATGCCACCATCTCAGCTGCTCTATATATATAAAGAAAGAAATTCAA 206
QY 287 ATTGCAAAATGAGCCCACTGTGAGGAGCACTAATGTGTCAATGTCACCTATTCGGCGCAC 346
Db 207 ACTAGAAAGTGAGCCTACTGTGAGGGGCTGATTGTGCCATGTCACCTAATTGAGCAAC 266
QY 347 CATGCTTACTACATTTCTCCGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTAT 406
Db 267 TATCATTAACAACATTTCTGTGACCAAGTATCAGACTGGAATTGGCCGCTCCCTATGCTCAT 326
QY 407 TCTCTTCAATTTCTGTACTTCTTGACGGGCTCATGATGCTATAGTCTCTCAATGCTTA 466

Db 327 CCTCTCTTCAAGTTCGTACTTTGTTAGCAGCCTCATATGCTATGGTCCCGAATGCTA 386
QY 467 TGTCTCTGCTGTGGSCACGCTTCGTAGATGGAATTTGGTATGGCTTGCGTCAAGCTTGT 526
Db 387 TGTACTACTGCTGGSCACGCTTCATATAATGGAATTCGCTGTCGGCTTGCGTCACTCTTGT 446
QY 527 GCCTTTGTACATTTCAGAATAAGCCCCCTTCGAGATTTAGAGGTTTGTCTGAATACACTACC 586
Db 447 TCCTCTGTACATTTCAGAACAAGCTCTCCAGAGATTAAGAGGTTTGTCTGAATACACTGCC 506
QY 587 ACAATTCAGTGGATCAGAGAGAAATGTTCTTGTCAATCTGATGGTGTGGGATGCTCCT 646
Db 507 ACAGTTCAGTGGGTCAAGGGGGAGTGTCTTGTCTGATCTGATGGTGTGGGATGCTACT 566
QY 647 GTGECATCACCCGATTTGGAATATATGCTTGTGTGCTGCGCATACCTTCATTTGTTCTT 706
Db 567 GCTGCATCACCTGATTTGAGGATATGCTTGGGGTCTCGCGTTACCTTCATTTGTTT 626
QY 707 CTTTGGTTGACAATATTTTATCTTCTCGAATCTCCAAGATGGCTCGTTAGCAAGGTCG 766
Db 627 CTTTGGTTGACAATATTTTATCTGCTCGAATCAACAAGTGGCTGTCAAGCAAGGGCG 686
QY 767 GATGGCAGAGGCAAAAAGGTGTTCAAAAAGTTACGGGGGAAAGACGATGTCTCAGCTGA 826
Db 687 AATGGCAGAGGCAAAAGAGGTGTTCAAAAAGTTACCTGGGAAGGATGACGTCAAGGCGA 746
QY 827 ATGTCCCTTCTTCTCGAAGGTTGAGGTTGAGAGACACTTCCATTTGAAGATGACAT 886
Db 747 AATGGCTCTTCTTGTGAAGGTTGAGGTCGAGGGGATACCTTCATTTGAAGATGACAT 806
QY 887 CATTTGACCTGCCACCGAGGAGCCGATGATCTTGTACTGACGGGTGATTAAGGAACAAT 946
Db 807 CATTTGCTCTGCTATCGAACCAAGCTGATGACCAAGTCGCTGATGGTATAGAGAAGCAAT 866
QY 947 CACACTTATGGGCTGAAGAGGCCAGTCATGATGCTCGACCTTCTAAGGAACCCAT 1006
Db 867 AATACTATTTGGGCTGAAGGGCCCATCATGATGATCGGCTTCCAAAGGGTCCAG 926
QY 1007 CATGCTTGAAGTGTCTTCTTCTGATCTCGTCAATGGAGCATGGTGAACCAAGTGT 1066
Db 927 CATGCTTGAAGTGTCTTCTATCGATCTCGTCAATGGAGCATGGTGAACCAAAATTT 986
QY 1067 ACCCCTTATGGATCCGATTTGACACTTTTGGTACTGTCCATGAGAAATATGCTCAAGC 1126
Db 987 ACCCCTTACGGATCTTATGTCAAGCTTTTGGAGTGTCCATGAGAACATCTCTCTGC 1046
QY 1127 TGAAGGAAGTATGAGGAGCACTGTTTCCAAACTTTGGAAGTATGTTCACTGTCAAGA 1186
Db 1047 TGAAGGAAGTATGAGGAGAAATAGTCTTCTTAAATTTGACAGTATGATCAATCAAGA 1106
QY 1187 TCAGCATGCCAAAAATGACAGTGGGATGAAGAAATCTTCATAGGATGACGAGAGTA 1246
Db 1107 TCAGCAACCCCAAGACCGAGAGTGGGATGAGAGAACTCCATAGGATATAGAGAGTA 1166
QY 1247 CGCATCTGATGGTGCAGAGAGTGAATATGAGGACAAATCTCCATAGCCCATTTGCTCCAG 1306
Db 1167 TGCAACTGATGGTGTCTGCAAGTGAATGAAAACAATGTACACAAGCCCTGTGCTCCCG 1226
QY 1307 GCAGGCAACAGGTGCGGAAGGGAAGGACATTTGTGCACATGGTCAACCGTGAAGTCTTT 1366
Db 1227 ACAGAATACAAAGTCCGAAGGGAAGGGCATTTGCACATCAAGGGCACCGTGAAGTCTTT 1286
QY 1367 GAGCATGAGAAGGCAAACTCTTAGGGGAGGGTGAAGATGGTGTGAGCAGACTGATAT 1426
Db 1287 GGGTTTGAAGAAGAAGCCTGTCTCGATGAGGGTGTGAGGCAACAACGACACTGGCAT 1346
QY 1427 CGGTGGGGATGGCAGCTTCTTGAATATGTCAAGAAAGGTGAAGATGTTAGAAA 1486
Db 1347 TGGTGGCGGATGGCAGCTGCGCTGGAATGTCAAGCAGAGAAGGTGAGAGCGGTAAAGAA 1406
QY 1487 GGAAGTGTGTTCAAAAAGATCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAAAGGG 1546
Db 1407 GGAAGGAGTTCAAAAAGATCTACTTACACCAAGAGGAGTGTGCTGCTCAAGAAAGGGG 1466

QY 1547 CTCATTTGTTCACTTCCCGGTGTGGCGATGTTTGAAGGTAGTGAAGTTGTATCATGC 1606
Db 1467 ATCTGTGTCTCACTTCTCGGTGAGGTGATGCTCTGAAGGTGTGATGATCATCATGC 1526
QY 1607 TGCTGCTTTAGTACTGACTGACACTTTTCTCAAAAGGTCCTGTGAACCAAGCATGTC 1666
Db 1527 TGCTGCTGTGGTCAAGCATGACGCTCTTACCCGAGGATATTTACAGAACGCGCATGCG 1586
QY 1667 AGATGCTCCATGGTTCACCCCATCTGAGGTAGCTGCCAAAGGTTACGTTGGAAGAATTT 1726
Db 1587 CGGTCCAGCTACATGACACCCATCAGAGGAGCTGCCAAAGTTCCAAGCTGAGAGACCT 1646
QY 1727 GTTTGAACCTGAGTGAAGCGGTGCCCTGTAGTCGGTGTGGAATTCAGATCCTTCAACA 1786
Db 1647 GTTTGAACCTGGTGTGAGACGTCGCCCTGTGTGTTGTTATGGAATTCAGATCCTTCAACA 1706
QY 1787 GTTTGCTGAATAAACGGTGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGGTGT 1846
Db 1707 GTTCGGGGAATTAATGGGTTCTCTACTACATCCCGCAAAATTAAGAGCAAGCTGTGT 1766
QY 1847 GGCAGTATCTTTCCAAATTTGGTCTCAGCTGGCATCAGCATCCATCTTGAATCAGTTC 1906
Db 1767 GGCAGTCTTATTTCCAAATCTTGTGTCAGTTGCGCATCAGCATCCATCTTAAATGATTC 1826
QY 1907 TCTCACTACCTTACTAATGCTTCTTGCAATGGCTTTGCCATGCTGTTATGATCTTTC 1966
Db 1827 TGTCACTGCCCTACTATGCTCCCTAGCATGGTTAGCCATGAGACTTAAGACGTATC 1886
QY 1967 CGGAAGAAGTTTGTCTGCTAGGCAACAATTCCAATCTTGATAGCATCTAGTTATCCT 2026
Db 1887 TGAAGAAGGTTTCTGCTGCTAGCAACAATTCAGTCTTGATAGCATCTTAAATGTCTCT 1946
QY 2027 GGTGTGTCCATCTAATTTGATTTGGGTACACTAGCCCATGCTTGTCTTCCACCGTCAG 2086
Db 1947 GGTGTGTCTAATGTTATTCAGATTTGGGTACAGTGTCCATGCTGTGTCTTCCACAATTAG 2006
QY 2087 TGTATCTCTACTTCTGCTGCTTCTGTTAAGGATTTGTGCCATCCCAACATTTATG 2146
Db 2007 TGTATCACCCTACTTATGCTGCTTCAAAATGGGCTTTGGCCCATTCCAACATTTATG 2066
QY 2147 TGCAGAGATCTTCCAAACGAGGTTGCTGCTCTGTATTTGCCATTTGCTTACAT 2206
Db 2067 TGCAGAGTTTTCCAAACGAGGTTGCTGCAATGCTGATGCCATCTGCTGATAT 2126
QY 2207 CTGATCGAGATATCATGCTCACCTACAGCTTCTCTGTGATGCTGAATGCTATTTGACT 2266
Db 2127 TTGGTTGAGACATTAATGTCACATACAGTCTTCTGTGATGCTGAATGCTATTTGCT 2186
QY 2267 GCGGGGTCTTTCAGCATATATGACAGTCTGATTTGATTTCTTGTGTCTTCTCT 2326
Db 2187 AGAAGGCGTGTTTGGCATATACGACAGTGTGCTGCTGCTTGTGTGTCTATCT 2246
QY 2327 TAAAGTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTAACGAATCTTTGCAAGTTGG 2386
Db 2247 TAAAGTCCGAGACAAAGGGGATGCCCTTGAGGTCATCACTGAGTCTTTGCAAGTTGG 2306
QY 2387 TCGAAGCAAGCGCTGCAAAAGCCTAATTTCTTTGGTACCTTTGTGCAACTATTGCA 2446
Db 2307 TCGAAGCAAGCGGTTGCAAAAGCCTAGTCTTCCGCAACTTTGCTTGAACCTTTTGCA 2366
QY 2447 CTGTAAGTTAGAACTTGAAGGGGTTTCAACCAAGAGCTCGAGAAATTACTTTGGAATTTG 2506
Db 2367 CCAT-----GCAACTTGAAGGATCTCACCAAGAGTTTG-----GAGGTTTTGATC 2415
QY 2507 TGTAAATGTTAAGGGAACGAACATCTGCTCATGCTCTCAAAACGGTAAAAAAGAGTCCCT 2566
Db 2416 TGTAAATGAGGGAACGACAGACATGTGCCATTTCTCTAAAGCGACACTGATCAATGGCA 2475
QY 2567 CAATGGCAATAGAGTGTGTAAGTTGCAATGTCATTTACCATATG-----TT 2615
Db 2476 AAGAACAGCTTTTGGGTAGCGAAAT--TCGCTGTCATTTACCATCAGTTATGCTTATTT 2533

QY 2616 TTACCTATTGTACTGTATATATAGTCAAGCTATTCAACCGCTGGTGTGCTAGAAATCT 2675
Db 2534 GTACTTATTGTACTGTATATAGTATGTGCTGTTAATAGCTGCTTGTGTGGAGATGT 2593
QY 2676 TT 2677
Db 2594 TT 2595
RESULT 3
HVU534445
LOCUS HVU534445 2614 bp mRNA linear PLN 14-FEB-2003
DEFINITION Hordeum vulgare mRNA for hexose transporter (stpl gene).
ACCESSION AJ534445
VERSION AJ534445.1 GI:26986185
KEYWORDS hexose transporter; stpl gene.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Hordeum.
REFERENCE
1 Weschke,W., Panitz,R., Gubatz,S., Wang,Q., Radchuk,R., Weber,H. and
AUTHORS Wobus,U.
TITLE The role of invertases and hexose transporters in controlling sugar
ratios in maternal and filial tissues of barley caryopses during
early development
JOURNAL Plant J. 33 (2), 395-411 (2003)
MEDLINE 22424051
PUBMED 12535352
REFERENCE 2 (bases 1 to 2614)
AUTHORS Radchuk,R.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene
Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3,
Gatersleben, D-06466, GERMANY
FEATURES
source
1. 2614
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/db_xref="taxon:112509"
/tissue_type="developing caryopsis"
/clone_id="lambda ZAP2"
/dev_stage="0-15 days after flowering"
1. 2614
/gene="stpl"
83. 2314
/gene="stpl"
/codon_start=1
/product="hexose transporter"
/protein_id="CAD58958.1"
/db_xref="GI:26986186"
/translation="MSGAVLVVAIAISGNLQGMNATAGAVLYIKKESLETQPL
IEGLIVAMSLIGATVITTFSGAVADAVGRPLIASVLYFVSGVLMWAPVYVLL
ARLDGFGIGLAVTLVLYISETAPTDIRGLNTLPQFSGSGMFLSYCMVFTMSLMP
OPDWRIMGLVLSIPSLMYFALTVEFYLPESPRMVSKGRMAEAKRVLQRLGRBEDVSGE
MALLVBEGLVGKDTHEEYIIIPDDELADDLAPDQEKLYGAEEGVSWIARPVKXG
GOSALGSLGLMSRHGSWYSGKSLVDPLVTLFGSVHEKMBEVMGSMRSTLFPNGSM
FSVAEQOQAKADWDASHRDEDDYASDHGADDIEDNINSPISRQATSVGEKEIAAPH
GSIMGVSSSMGGDAVSSMIGGQWLAWKTEREGADGQKEGFORIYLHEEVS
GDRGSIILMPGGDIPGGEYIQAALVSPALYSKDLIEQLAGPAMVHPSBAVAKG
TKMAELFERGVKHALFVIGIQLIQFAGINGVLYTPQILEQAVGILSLNIGLSS
SASILISALTLLMLPSIGIAMRMSGRFLSTIPVLIVAIILVAVVLDVGT
MVHAALSTISIVVFCFVWGFGPIPNILCAEIPTSVRGICIAICALTFWIGDIIVT
YTLFVMLNAIGLAGVFGIYAVVCMIAFVFMKVPETKMPLEVITEFFSVGAKQKE
ATD"
BASE COUNT 563 a 586 c 755 g 708 t 2 others
ORIGIN
Query Match 38.7%; Score 1075.4; DB 8; Length 2614;
Best Local Similarity 69.2%; Pred. No. 6.6e-272;
Matches 1549; Conservative 0; Mismatches 666; Indels 24; Gaps 5;

QY 165 TTGAGAGACAAGATGTGGGGGGCTGTCTTGTGCGCATAGTCGCCCTCCATCGGCATCTA 224
Db 74 TCGGGCGGCATGTCTGCGGGGGCGCGGTGTGTGCGCATAGCGGGCGTCCATCGGGAACCTG 133
QY 225 TTGCAGGGGTGGGACATGCCACCATCGCAGCTGTCTGTATATATAAGAAAGAAATTT 284
Db 134 CTGCAGGGGTGGGATATATGCCAATTGCGAGGTGCGGTCTGTACATATAAGAAAGAAATTT 193
QY 285 CAATTGCAAAATGAGCCCACTGTGAGAGGAGCTAATGTGTCAATGTCACTTATCGCGCC 344
Db 194 AGCTTGAGAGACCCAGCCCTTGATCGAGGCGCTCATCGTGGCCATGTGCTCATCGAGCG 253
QY 345 ACCATCGTTACTACATCTCCGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTT 404
Db 254 ACGGTTATCACGACGTTCTCCGGGCGGTGCTGACGCTGTTGTAGCGCGCCCTGCTG 313
QY 405 ATTCTCTTCAATTCTGACTTCTTACGCGGCTCATCATGCTATGGTCTCTAATGTC 464
Db 314 ATCGCCTGCTGTCTCTACTTGTCACTGCGCTGTGATGCTGTGGCGCCCAACGTC 373
QY 465 TATGCTCTGCTGTGGACGCTTCGTAGATGATTTGGTATTTGGCTTGGCTTCACGCTT 524
Db 374 TATGTGTGCTCTTGGCCAGGCTCATCGACGGGTTCGGTATCGGTTGGCAGTCAACCTT 433
QY 525 GNGCCTTGTACATTCAGAAATAGCCCTTCGGAGATTAGAGTTGCTGAATACACTA 584
Db 434 GTCCCTGTACATTCAGAAACCGCCCGACTGTACATTAGAGGCTGTGACACGCTG 493
QY 585 CCACAATTCAATGATCAGAGAGGAATGTTCTGTACATCTCATGATGTTGGATGTC 644
Db 494 CCGCAGTTCAATGATGATCAGAGAGGATGTTCTTCTTACTGCATGATGTTTACATGTC 553
QY 645 CTGTGCCATCACCCGATTTGGAAGATTATGCTTGTGCTCGCATACCTTCATTTGTC 704
Db 554 CTCATGCCGACGCTGATCGAAGATCATGCTTGGGTTTGTGATCCCGCTGATG 613
QY 705 TTCTTTGGTTGACATATTTTATCTTCTGATCTCCAAGATGGCTCGTTAGCAAGGT 764
Db 614 TATTTTGCATGATCTGCTTTTATTTGCCCCGAGTGCACAGATGGCTGTGACCAAGGA 673
QY 765 CGGATGCGAGAGCAAAAAGGTGTTGCAAAAGTTACGGGGGAAAGACGATGTTCAAGT 824
Db 674 AGAATGCTGAGGCCAAGCAGATATTGCMAAGACTCGGGGAAAGGAGATGTCAGGA 733
QY 825 GAATGTCCCTTCTTGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGATG 884
Db 734 GAGATGGCCCTTCTTGTGAAGATTGGGTGGAAGAAAGACACACATTTTGAAGAAATG 793
QY 885 ATCATTTGACCTGCAAGGAGGAGCGCATGATCTTTACTGACGATGATAGGAACA 944
Db 794 ATAATTGACCTGATGATGAGCTGCTGATGACGGGCTGGCT---CCAGATCAAGAGAAG 850
QY 945 ATCACACTTATGGGCTGAAGAGGCGCAGTCATGATGCTCGACCTTCTAG----- 998
Db 851 TTGAAACTCTATGAGCTGAAGAGGGGTATCTTGAATGCCCGCTCTGTAGSSGTGGC 910
QY 999 GGAATCATCATGCTTGAAGTGTCTTCTTGTGATCTCTCATTTGGAGCATGCTGAAC 1058
Db 911 GGCCTAAAGTCACTTGAAGCGCTTGGGCTCATGCTCTGTCATGAGATGTTAGT 970
QY 1059 CAGAGTGTACCCCTTATGATCCGATTGTGACACTTTTGTAGTGTCCATGAGAAATATG 1118
Db 971 CAGGTAATCTCTTGTGACCCCTTGTCTCTTCCGAAAGTGTCCATGAGAAATG 1030
QY 1119 CTTCAAGCTGAGGAAGTATGAGAGACACATGTTTCCAACTTTGGAAGTATGTTCACT 1178
Db 1031 CTTGAGGTATATGGGAGATGCGGAGACCTTGTTCCTAATTTGGCAGCATGTTAGT 1090
QY 1179 GTCAAGATCAGCATGCCAAAATGAGCAGTGGATGAAGAGATCTTCATAGGATGAC 1238
Db 1091 GTGGCTGAACAGCAGCAGGCTAAAGCTGACTGGGATGCTGAGA---GTCAATAGGATGAT 1147

QY	1239	GAGAGTACGCATCTGATGTCAGAGAGTGACTATGAGACAATCTCCATAGCCCATG	1298
Db	1148	GAAGATTATGATCGATCATGTGCTGATGACATGAGATTAACCTCAATAGCCGCTT	1207
QY	1299	CTGTCCAGGCAGCAACAGTCCGGAAGGAGACATTTGCAACCATGTCACCGTGA	1358
Db	1208	ATTCTCTGTCAGCGACACAGCGGTGAGGGTAAGAGATTTCTGCACCTCATGAGACATA	1267
QY	1359	AGTCTTTGAGCATGAGAAGCAAAACCCTTAAAGGAGGTGAGATGTTGAGCAGC	1418
Db	1268	ATGGGT-----GGTGTGAAAGTAGTAGCATGCAAGGAGGGACGACAGTAGCAGC	1318
QY	1419	ACTGATATCGGTGGGGGATGGCAGCTTGCTTGAAATGTCAGAGAAGGAAGGTGAAAT	1478
Db	1319	ATGGGCATTGTGTGGGGGTGGCAGTTAGCTTGGAATGGACTGAGAAGAGGTGCAGAT	1378
QY	1479	GGTAGAAAGGAAGGTGGTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCTTGCTCA	1538
Db	1379	GGGCAAAAAGAGGTGGCTTCCAGCGTATTACTTGCATGAGAGAGGGTGTCTCAGGTGAT	1438
QY	1539	AGAGGGGCTCAATTGTTTCACTTCCCGGTGTGTCGATGTTTGGGGTATGAGTTT	1598
Db	1439	CGAGAGGCTCTATATTGTATTATGCCAGAGAGT---GATAATCCTCTGTGTGTGATAT	1495
QY	1599	GTACATGCTGCTGCTTTAGTAAGTCAGTCAACACTTTTCTCAAGGGCTTGTCTGAACA	1658
Db	1496	ATCCAGGACGCCGCTCTAGTGAAGCAACCTGCTCTTTACTCGAAGACCTGATAGACAG	1555
QY	1659	CGCATGTCAGATGCTGCCATGGTTCAACCCATCTGAAGTAGCTGCCAAGGTTCAAGTTGG	1718
Db	1556	CAGCTGTGCTGCCAGCCATGGTATCATCCATCCGAGCGAGTTGCCAAGGTTACAAAGTGG	1615
QY	1719	AAAGATTGTTGAACCTGAGTGAAGCGGTGCCCCGTGTACTCGGTGTTGGAATTGACATC	1778
Db	1616	GCGAAGTATTGAACTGGAGTGAAGCATGCACTGTTGTTGGCAATTGGAATTACAGATC	1675
QY	1779	CTTCAACAGTTTGCTGGAATAAAGGTGTTCTGTACTATACCCACAAATTCTTGAGCAA	1838
Db	1676	CTGCAGCAGTTTGGCGGTATCAACGAGTCTGTATTAACAAGCTCAGATATTGAGCAA	1735
QY	1839	GCTGTGTGGCAGTTATTTCTTCCAAATTTGTTGTTCTAGCTCGGCATCAGCATCCAATCTTG	1898
Db	1736	GCAGGTGTTGGATTCTTTATCAAAACATTGAGACTAAGCTTCTCTGATCTATTTCTT	1795
QY	1899	ATCAGTTCTCTCACTACCTTAATAAGCTTCTTGCAATGGCTTTGCCATGCTGCTATG	1958
Db	1796	ATTAGTGCCCTTGACAACCTGTGTGATGCTTCCAGCAGTTGGCATCGCCATGAGACTATG	1855
QY	1959	GATCTTCCGGAAGAGTTTGTCTGCTAGGCAATTCCAATCTTGATAGCATCTCTA	2018
Db	1856	GATATGTCAGGAAGAGTTTCTTCTCCTCTCAACATCTCTGTTCTTGATAGTGGCACTA	1915
QY	2019	GTTATCTGTGTGTCCAACTTAATTGATTTGGGTACACTAGCCCATGCTTGTCTCC	2078
Db	1916	GCTATCTGTGTTTGGTGAATGTTCTGAGATGGGAACCATGTTGCACGCCGCTCTCTCA	1975
QY	2079	ACCGTCAGTGTATCTCTACTTCTGCTGCTTCTGTATGGAATTTGTTGCCATCCCAAC	2138
Db	1976	ACGATCAGTGTATAGTCTCTCTGCTTCTGTCTGATGGGTTTGGACCTATCCGAAT	2035
QY	2139	ATTTTATGTGACAGATCTTTCCAACCCAGGGTTCGNGGCTCTGTATTGCCATTGTGCC	2198
Db	2036	ATTCTCTGCGCGAGATTTTCCCACTCCGTCGNGGCATCTGCATTGCCATCTGGCGCA	2095
QY	2199	TTTACATTTCTGATCGAGATATCATCGTCAACCTACAGCCTTCTGTGATGCTGAATGCT	2258
Db	2096	CTAACCTTCTGATTTGTTGATCATCGTGACATACACTTCCCGTATGCTCAACGCC	2155
QY	2259	ATTGCACTGGCGGTGTTTCAACATATATGCAAGTCTATGCTTGATTTCTTGTGTTC	2318
Db	2156	ATCGCACTCGCTGGAGTTTTCGAATATATGCTGTCTGTTTGCATGATAGCCTTGTATTC	2215
QY	2319	GTCTTCTTAAGGTCCCTGAGACAAAGGGGATGCCCTTGAAGGTATTACCGAATCTTT	2378

Db 2216 GTCTACATGAAGGTCCCTGAGACAAAGGGCATGCCCCCTGAGGTCATCACCGAGTTCTTC 2275

QY 2379 GCA GTTGTGCGAAGCAAG 2397

Db 2276 TCGGTCGGGGCGAAGCAGG 2294

RESULT 4	AP005756/c	LOCUS	DEFINITION
	AP005756	137140 bp	DNA linear
	Oryza sativa (japonica cultivar-group)		chromosome 2 clone
	OSJNBd0035N08, ***	SEQUENCING IN PROGRESS ***.	

ACCESSION	AP005756
VERSION	AP005756.1
KEYWORDS	GI:23200615
SOURCE	HTG; HTGS_PHASE2.
ORGANISM	Oryza sativa (japonica cultivar-group)
	Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE	1
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE	<i>Oryza sativa</i> niponbare (GA3) genomic DNA, chromosome 2, BAC clone: OSJNBb0035N08

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Published Only in Database (2002)
2 (bases 1 to 137140)
Sasaki, T., Matsumoto, T. and Katayose, Y.
Direct Submission
Submitted (18-SEP-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasakienias.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>, Tel: 81-298-38-7441, Fax: 81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

- * NOTE: This is a 'working draft' sequence.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

FEATURES	Location/Qualifiers
source	1. 137140

```
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
```

BASE COUNT	ORIGIN
40995 a	27904 c 27671 g 40414 t 156 others

Query Match	34.1%;	Score 946.8;	DB 2;	Length 137140;
Best Local Similarity	69.9%;	Pred. No. 7.2e-238;		
Matches 1520;	Conservative 0;	Mismatches 407;	Indels 249;	Gaps 8;

Oy 814 ATGCTCAGGTGAATTGTCCCTTCTTCGAAAGGGTTGAGCGTTGAGAGACACTTCCA 873
 || | ||||| ||| | | ||||| ||||| ||||| ||||| |||||
Db 86221 ATTGCCAGGAGAATAATGGCTCTTCCTTGTAAGGTTTGAGGTTGAGAGCTGCACACTTCCA 86162

874 TTGAGAGTACATCTTGACCTGCCACCGAGCAGCCGATATCTTGTTACTGACGGTG 933

Db 86161 TTGAAGAGTACATCATCGACCCGGCTATAGAGCCAGCTGATGATGATGCG 86102

934 ATAGGACAATCACCTTATGGGCTGAGAAGGCCAGTCATGATTTGCTGACCTT 993

Db 86101 ATAGGACCAATAACACTGTACGGGCTGAAGAGGGCCAATCATGGATTGCTGCACCTT 86042

994 CTAGGGACCCCAICAGCTTGGAGTGTGCTTCTCTGGCATCTGTCATGGAGCAGG 1053

Db 86041 CCAAGGACCAGCATTCTTGGAAAGTGTGCTTCTCTTACATCTCGTCATGGCAGCATGG 85982
QY 1054 TGAACCCAGAGTATACCCCTTATGGATCCGATTGTGACATTTTGGTAGTTCATGAGA 1113
Db 85981 TGAACCCAGAGCGTCCACTTATGGATCCTATAGTACGCTTTTCGGCAGTGTCCATGAGA 85922
QY 1114 ATATGCCCTCAAGCTGGAAGAGTATGAGAGCAGCATGTGTTCCAAACTTTGGAATATGT 1173
Db 85921 ACATGCCCTCATGTGGAAGAGCATGCGAGTACATGTGTTCCAACTTTGGTAGTATGT 85862
QY 1174 TCAGTGTACAGATCAGCATGCGCAAAATGAGCAGTGGGATGAAGAAATCTCATAGGG 1233
Db 85861 TTAGTGTACAGATCAGCACCCCAAGTTGATCAATGGATGAGAGAACCTTCATAGGG 85802
QY 1234 ATGACGAGAGTACCGCATCTGATGGTGCAGAGAGTGACTATGAGGACAATCTCATAGCC 1293
Db 85801 ATGATGAGGAATACCGCTCTGATGGTGTGCTGGAGGTGACTATGAAGACAATGTCCACAGCC 85742
QY 1294 CATGTGCTGCCAGGCAAGGCAACAGGTGCGGAAGGAGACATTTGTGCAACCATGTGACC 1353
Db 85741 CGTGTCTGCCGACAGACCAAGCGCAGAAAGGAGACATTTGCACACCATGCTCATC 85682
QY 1354 GTGGAAGTGTCTTGAAGCATGAGAAGGCAAAACCTCTTAAAGGAGGGTGAAGTGTGTA 1413
Db 85681 GTGGAAGTGTCTGAGTATGAGAAGAAGAGCCTCTTGAAGAAGGGTGGCGAGGGGTGA 85622
QY 1414 GCAGCACTGATATCGGTGGGGGATGGCAGCTTGTGGAATGCTCAGAGAAGAGTGTG 1473
Db 85621 GCAGCACTGCATTTGGTGGGGGATGGCAGCTTGCATGGAATGCTCAGAGCGAGAGGCG 85562
QY 1474 AGAATGTAGAAAGAGAGTGTGTTCAAAAGAGTCTTGCACCAAGAGGAGTTCTCTG 1533
Db 85561 AGGATGTAGAAAGAGAGAGATTTAAAAGAGTCTTGCACCAAGAGAGAGTGCAG 85502
QY 1534 GCTCAAGAAAGGGGCTCAATTGTTTCACTTCCCGGTGGTGGCATGTTTGTGAGGTAAGT 1593
Db 85501 GATCAAGAAAGGGGCTCAGTTATTTCACTTCTGTGGAGGGGAGCGCTCTGAAGCAGCG 85442
QY 1594 AGTTGTACATGCTGCTCTTAAAGTCAAGTCAAGCATTTTCTCAAGGGTCTGCTG 1653
Db 85441 AATTCTACATGCTGCGCTTGGTAAAGCCAACAGCAGCATTTACTCCAAGATATCATTG 85382
QY 1654 AACCAAGCATGTAGATGCTGCCATGGTTCACCCCATGTAGAGTAGCTGCCAAGGTTAC 1713
Db 85381 AACAGCGTATGTCCGGTCCAGCCATGATTCATTCATCAGAGCAGCTGCCAAGGTTCAA 85322
QY 1714 GTTGGAAGATTTGTTGAACCTGGAGTGAGGCGCTGCTTAAAGTGTGTAATTC 1773
Db 85321 GCTGGAAGATTTGTTGAACCTGGAGTGAGGCGCTGCTTAAAGTGTGTAATTC 85262
QY 1774 AGATCCTTCAA----- 1784
Db 85261 AAATCCTTCAACAGGTAGAGATGATACCAATTTCAATAATCTTTTGTACTGTCCAACA 85202
QY 1785 ----- 1784
Db 85201 TTGACGGTGTTAAGATTTGTACTATTAAGAATTCACTAATTCCTTTATGTTCTTTC 85142
QY 1785 -----CAGTTGCTGAATAAACGCGTGTCTGTACTATACCCACAATTTCTGAGCAA 1838
Db 85141 CTGTGCAAGTTTGTGGAATAAATGGGGTCTCTATATACACTCCACAATATCTTGAACAA 85082
QY 1839 GCTGCTGTGCAAGTTATTTTCCAAATTTGCTCTCAGCTCGGCATCAGCATCCATCTTG 1898
Db 85081 GCGGGGGGTGCAAGTTCTCTTCCAAATCTTGCCCTCAGTTCAAGCATCAGCTTCCATCTTG 85022
QY 1899 ATCAGTTCTCTCACTACCTTAATAATGCTTCTGCTGATGGCTTTGCCATGCTGCTATG 1958
Db 85021 ATCAGTTCTCTGACCACCTACTGATGCTTCTTAGCATGGTTTAGCCATGAGACTTATG 84962
QY 1959 GATCTTTCCGGAAGA----- 1973
Db 84961 GACATCTCTGGAAGAAGTACTAATTGTTTCTTGTGTTGATCATCTTAGTACTACCA 84902

QY 1974 ----- 1973
Db 84901 AATATACCAAGAAATGACACCAAGTTAGTATTAATTATCATCATATGTTCTT 84842
QY 1974 ---AGTTTGTGCTGCTAGGCAACAATTCAACTTGTATGACATCTCTAGTTA 2030
Db 84841 TTTAGTTTCTGCTTCTGGGCACAATTCAGTCTTGATAGCATCTTAGTTGTCTT 84782
QY 2031 GTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTGCTCCACCGTCA 2090
Db 84781 GTGTCCAATGTTATGACCTGGGTACAGTGGCCACCGCAGCTCTCCACAATCAGCGTC 84722
QY 2091 ATGCTACTTCTGCTGCTTCTGTTATGGGATTTGGTCCCATCCCAATTTATGTGCA 2150
Db 84721 ATCATCTACTTCTGCTGCTTCTGTCATGGGATTCGGTCCGATCCCAACATCTGTGCG 84662
QY 2151 GAGATCTTCCAAACGAGGTTCTGCGCTCTGATTTGCCATTTGTGCTTACATTTCTG 2210
Db 84661 GAGATCTTCCCACTAGGCTCCGCGCATCTGCATTTGCCATCTGCGCTGACATTTCTG 84602
QY 2211 ATCGAGATATCATCTGTCACCTACAGCTTCTGTGATGCTGAATGCTAATTGGA 2270
Db 84601 ATTGTGATATCATAGTACCTACAGCTCTCTGTGATGCTGAATGCTAATGCTGGA 84542
QY 2271 GGTGTTTACGATATATGATGATGCTGATGCTGATTTCTTGTGCTGCTTCTTAA 2330
Db 84541 GGTGCTTGTGATATACGAGTGTGCTGATTTGCTGATTTGCTGCTTCTTAA 84482
QY 2331 GTCCCTGAGACAAAGGGATGCCCCCTGAGGTTATTAACGAATCTTGCAGTTGTCG 2390
Db 84481 GTCCCGAGACGAAGGGCATGCGCTGAGTATCATCACGAGTCTTCTGCTGTGTCGA 84422
QY 2391 AAGCAA---GCGGCTGCAAAAGCCTAATTTCTTGTGAC---CTTGTGTGCAACTATT 2443
Db 84421 AAGCAATGCAAGGCTACAAAGCCTAATTTCTTCAAGCAACTCTGCTTCAATCTTATATA 84362
QY 2444 GCACTGTATGTAAGAACTTGAAGGGGTTTCAACCAAGAGCTCGGAGAAATTA 2503
Db 84361 ACACTGTAAATGGAATCTGCAA--GATCAACCATGAGGCTGAGAGCTTTTGA 84304
QY 2504 TTGTGTAATGTTAAGGGAACGAACATCTGCTCATGCTCTCAACGTAATAAAGAGTC 2563
Db 84303 GT-----AAATGAGAGAAGAGATGCTAATTCATGCTCTTATGCAAGGAAAAAGATT 84250
QY 2564 CCTCAATGGCAATAGAGTCTTAAAGTGTCAATGTCATTTACATATGTTTACCTAT 2623
Db 84249 CATTACTGGCAGAGGCGAGCTGTAAAGTAGAATAATTTTCAATGTTATCATGTCACCTTCT 84190
QY 2624 TTGTACTGTATTA-----TAAGTCAAGCTATTTCAACGCTGTTGTGCTAGAATC 2674
Db 84189 TGTATTATTAGTACTGTGTTAGGTAATGTTTATTTATCATCAGTATGCTGCTAGGATG 84130
QY 2675 TTTAGACAAGATGATATGATCTGATCTGATGTTATATATTTCAAAATCTCAAAATAAG 2734
Db 84129 TTCAAAAATAAGCT--TGCTAATCTGATTTTGAGCTACTATGTTGGGTTCAATAATAAG 84072
QY 2735 AAAATATGTTTCTCA 2750
Db 84071 TAAATTTTCTGTGCA 84056

RESULT 5
AR208568 2601 bp DNA linear PAT 20-JUN-2002
LOCUS AR208568
DEFINITION Sequence 7 from patent US 6383776.
ACCESSION AR208568
VERSION AR208568.1 GI:21509755
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2601)

AUTHORS Allen, S.M., Hiltz, W.D., Kinney, A.J. and Tingey, S.V.

TITLE Plant sugar transport proteins

JOURNAL Patent: US 6383776-A 7 07-MAY-2002;

FEATURES Location/Qualifiers

Source 1. 2601

BASE COUNT 648 a 543 c 644 g 766 t

ORIGIN

Query Match 33.9%; Score 940.2; DB 6; Length 2601;

Best Local Similarity 66.0%; Pred. No. 2.9e-236;

Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

```
QY 177 ATGTCGGGGGCTGTTCTTGTCCCAAGTCCCTCCATCGGCAATCTATTGCAAGGGTGG 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ATGAAGGTGCGGCTCTGTTGCTAATTGCCGCTCCATTGGTAATTTCCTCCAAGATGG 234

QY 237 GACAAATGCCACCATCGCAGCTGCTGTTCTGTATATAAAGAAGAAATTCAATTGCCAAAT 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 GATTAATGCTACCATCGCGGGGCTAATGTTACATTAGAAAGAACTTGCTTTGGAA-- 292

QY 297 GAGCCCACTGTGAGGAGCAATATGTGTCAATGTCACTTATCGGCGCCACCATCGTACT 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 ---CAACTATGAAAGGCTGTGTGGGCATGTCCCTGATGGAGCAACGTAATCAAC 348

QY 357 ACAATCTCCGGGCATTATCAGACTCGATTGGCCGACGCCCTATGCTTATTCTCTCA 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 ACAATGCTGTGCTCTATAGCGAATGGCTCGGTCCGCGACCATGATGATAATCTCATCT 408

QY 417 ATTCTGTACTTCTTCAGCGGCTCATCATGCTATGCTCTCTAATGTCTATGTCCTGCTG 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 GTGCTCTATTCTTGCGGTGTTGGTGATGCTGTGGTCCCAAAATGTGTATGTGTGGC 468

QY 477 TTGGCAGCGCTTCGTAGATGATTGGTATTTGGCTTGCTGTCAAGCTTGCTTGTAC 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 TTGGCGAGGCTACTTGATGATTGGGATTTGGCTTGTGTGACTCTTGTCCCGGTCTAT 528

QY 537 ATTTCAGAATATGCCCCCTTCGAGATTAGAGGTTTGTGTAATACACTCACAAATCACT 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 AATACTGAAACGGCGCGCTGTGAATAAGGGGTCGTTGAATACGCTTCTCAGTTCACT 588

QY 597 GGAATCAGAGGAATGTTCTGTCACTCATGCTGTTGGGATGTCCTGTGCGCACTCA 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 GGCTCTGAGGAATGTTTGTGCTACTGATGTTTGGCATGTCATTGATCCCGG 648

QY 657 CCCGATTGGAGATATATGCTGTGTGCTCGGATACCTTCATTGTTCTTGTGTTG 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 CCTAGCTGAGGCTCATGCTTGGGGTCTGTCTATTCTTCTCTTGTATTGTGACTG 708

QY 717 ACAATATTTTATCTTCTGTAATCTCCAAGATGGCTCGTTAGCAAAAGGTCGGATGGCAAG 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 ACCATTTTCTTCTGCCCCGAGTCTCTCGGTGGCTGCTCAGCAAAAGGAAGATGCTGAG 768

QY 777 GCMAAAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGCTCAAGTGAATTGCCCTT 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 GCTAAGAAAGGTGCTCCAAGATTGCGGAAAGGAGATGTCTCAGCGAGATGGCATTG 828

QY 837 CTTCCTGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGATACATCATTTGCACT 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 829 CTGGTTGAAGGTCTCGGATTTGGGGGTGATACATCATCGAAGAGTACATAATTGCCCT 888

QY 897 GCGACCGAGGACCGCATGCTTGTACTGACGGTGAAGAAACAATCACACTTAT 956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 889 GCTGACGATGTGCTGATGTGTGAACATGCAACAGAGAAAGATAAATTGATATAT 948

QY 957 GGGCTGAAGAAGGCCAGTCAATGATGCTGACCTTCTAAGGGAACCATCATGCTTGA 1016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 949 GGATCCCAAGCAGGCTTCTTGGTTATCAAAACCCTGCTACTGACAGAGTTCTATTGG- 1007

QY 1017 AGTGTCTTCTCTTGCACTCTGCTCATGGAGCATGTGTGAACAGAGTGTACCCCTTATG 1076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1008 -----CCTTGCTCAACCATGGAAGCATCATCAACAAAGCATGCCCTCATG 1056
```

```
QY 1077 GATCCGATGTGACACTTTTGTAGTGTCCATGAGATATATGCTCA-----AGCTGGA 1130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1057 GATCCTCTGTGTGACACTGTTTGGTAGCATTCATGAGAAAGCTCCCCGAGACAGAGCAAGA 1116

QY 1131 GGAAGTATGAGGAGCACATTTGTTCCAACTTTGGAAGTATGTTCAGTGTCAAGATCAG 1190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1117 GGAAGCATGCGAAGCACTCTGTTCCAAATTTTGGAAAGCATGTTCAAGCACTGCTGAACCG 1176

QY 1191 CATGCCAAAATGAGCAGTGGGATGAAGAGAAATCTCATAGGATGACGAGATACGCA 1250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1177 CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAGGAAACGTGAGACTACATG 1236

QY 1251 TCTGATGTGACAGAGTGACTATGAGAACAAATCTCATAGCCCATTTGCTGTCCAGGCA 1310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1237 TCAGATGCAACCCGTGGGGACTCCGATGATTAATTGCACAGTCCCTTAATCTCACGCCAA 1296

QY 1311 GCAACAGGTGCGAAGGGAAGACATTTGTGACCATGCTCACCGTGAAGTGTGAGC 1370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1297 ACAACAAGCCTTGAAGAAAGACTTACCTCCTCCTCCCATGGCAGTATCCTTGGCAGC 1356

QY 1371 ATGAGAAGCA--AACCTCTTAGGGGAGGGGTGAGATGCTGTGACGACGACTGATATC 1427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1357 ATGAGGCGTCAACAGTACTCATGCAAGGGTCAAGTGAAGCAAGGTGTGTATACAGTATT 1416

QY 1428 GGTGGGGGATGCGAGCTTGTCTGAAATGCTGAGAAAGGAAGGTGAGAAATGTAAGAA 1487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1417 GGTGTGGCTGGCAACTGGCATGGAATGACTGTATTA--AGGTGAGATGGAAGAAACAA 1473

QY 1488 GAAAGTGTTCAAAAGAGTCTACTTGACCAAGAGGAGGATTCCTGCTCAAGAGGGGC 1547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1474 CAAGAGGCTTAAAGAGATTATTATCATGAGAGAGGAGTTTCTGCATCTGCTGTGA 1533

QY 1548 TCAATTGTTCACTCCCGGTGTGGCGAGTGTTTTGAGGGTAGTAGTTGTATGATGCT 1607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1534 TCCATTGTATGATCCCGGTGAAGCGAA-----TTGTCCAGGCT 1575

QY 1608 GCTGCTTAGTAAATCAGTCAACACTTTTCTCAAGGCTTGTCTGAACACGATGTCA 1667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1576 GCTGCCTTGTGTAAGCAACCCGCTCTTACTTCAAGAGACTTATTGATGACACCCAGTT 1635

QY 1668 GATGCTGCCATGTTCAACCATCTGAGGTAGCTGCCAAAGGTTCAAGTTGGAAGATTG 1727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1636 GGGCTGCAATGTTTCAACCATCTGAGACAGCTTCAAGGGGCCAAGTTGGAAGAGCTCTT 1695

QY 1728 TTTGAACCTGAGTGAGGCGGTGCTGTAGTGGGTGTTGGAATTCAATCTTCAACAG 1787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1696 CTGGAACGAGGGTTAAGCATGATGTTGTGTGAGTTGGAATACAAATTACTTCAGCAG 1755

QY 1788 TTTGCTGAATAAAGGTGTTCTGTACTATAACCCCAAAATTTTGAGCAAGCTGTGTG 1847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1756 TTTTCAAGGATTAATGGGGTCTATATTACACACTCAATCTTGAAGAGCCGGTGT 1815

QY 1848 GCAGTTATTTCTTCCAATTTTGTCTCAGCTCGGCATCAGCATCCATCTGATCAGTTCT 1907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1816 GAAGTCTTCTTACAGATATAGCATTTGGCTCAGAGTCGGCATCATTCCTTATCAGTGT 1875

QY 1908 CTCACCTACTAATATGCTTCTTGCATTGGCTTGGCATGCTGTGATAGATCTTCC 1967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1876 TTCAACAACCTTCTTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGATGTTCA 1935

QY 1968 GGAAGAAGTTTGTGCTAGGACACAATTCGAATCTTGATAGCATCTCTAGTTATCTG 2027
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1936 GGCAGAAAGGCAAGTGTACTTACTACAATCCCGTGTGATGTGTGCTCATTTATTTG 1995

QY 2028 GTTGTCTCAATCTAATGATTGTTGGGTACACTAGCCCATGCTTGTCTCCACCGTCA 2087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1996 GTCAATTGAAGCTGTGTAATTTTGGCAATGTGCCCATGAGCAATCTCAACAGTATGC 2055

QY 2088 GTTATGCTACTTCTGCTGCTTGTATGGAATTTGGTCCCATCCCAACATTTTATGT 2147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2056 GTTGTGTTTATTTCTGCTGCTTGTGTATGGGTATGAGCAATTCAAACATCCTTGTG 2115

QY 2148 GCAGAGATCTTCCAAACAGGTTCTGTGCTCTGTATTGCCATTTGTGCTTATCATTC 2207
```

|||||
Db 2116 TCAGAGATTTTCCCACTAGGGTGGCTGGCTCTGCATGTCTATCTGTGCATTAAGTCTC 2175
QY 2208 TGGATCGAGATATCATCTGCACTACAGCCTTCCGTGATGCTGATGCTATGACTG 2267
|||||
Db 2176 TGGATTGGAGACATCATCATCACTACTCGCTGCTGCTGATGCTCGGCTCTTAGACTT 2235
QY 2268 GCGGGTGTTCAGCATATATGCACTGCTGATGCTGATTCCTTGTGTCTCTTCTT 2327
|||||
Db 2236 GGTGTGTATTCCCATTTACGACAGTTGTTGTTTCATCTGCTGATATTGTTGTTTG 2295
QY 2328 AAGGTCCCTGAGACAAAGGGGATGCCCTTGAGTTATTACCGAATTCCTTGCACTGCT 2387
|||||
Db 2296 AAGGTCCAGAAACAAAGGCGATGCCCTTGAAGTCACTCTGAATTCCTTCTGTTGGA 2355
QY 2388 GCGAAGCAAGCGCTGC 2404
|||||
Db 2356 GCAAGACAGGCTGCTTC 2372

RESULT 6
AY094465 2570 bp mRNA linear PLN 05-MAY-2002
LOCUS Arabidopsis thaliana AT4g35300/F23E12_140 mRNA, complete cds.
DEFINITION
ACCESSION AY094465
VERSION AY094465.1 GI:20453188
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Shimn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
TITLE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2570)

TITLE
JOURNAL Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory
(SIGAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
sequencing and annotation of the RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shimn,P., Chen,H.,
Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Chang,E.,
Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A.,
and Ecker,J.R.

Shimn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to

this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
source location/Qualifiers
1.2570

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="RAFL09-48-J19(R21968)"
/note="ecotype: Columbia"
209..2428
CDS
/note="putative sugar transporter protein"
/codon_start=1
/product="AT4g35300/F23E12_140"
/protein_id="AAM19835.1"
/db_xref="GI:20453189"

BASE COUNT 629 a 500 c 646 g 795 t
ORIGIN
Query Match 33.6%; Score 932.4; DB 8; Length 2570;
Best Local Similarity 65.8%; Pred. No. 3.3e-234;
Matches 1474; Conservative 0; Mismatches 736; Indels 30; Gaps 7;
168 GAGGACAGATGTCGGGGCTGTTCTTGTCGCATAGTCGCCCTCCATCGCAATCTATTG 227
200 GACGGAATTATGAGTGAGCTGCTGTTGCTTAATGCTGCTGCTGTTGCACTGTGTTA 259
228 CAGGGGTGGACAATGCCACCATCGAGCTGCTCTGTATATAAAGAGAAATTCAA 287
260 CAAGATGGATAACGACACTATTGCAGAGAGCTGTGTATACATAAAAAAGAGTTAAT 319
288 TTGCAAAATGAGCCCACTGTGAGGAGCTAATTGTCTCAATGTCACTTATCGGCCACC 347
320 TTGAGAGATTAATCCATCAGTGAAGGCTTAATGTGCGCATGTCACTTATGTGCTACT 379
348 ATCGTTACTACATTCCTCGGGCCATTATCAGACTCGATTGCGCGAGCCCTATGCTTAT 407
380 CTGATTACAACATGCTGTGAGGGGTAGCTGATGCTGTGCGCGCTCCATGCTTAATA 439
408 CTCTCTCAATTCCTGACTCTTCTCAGCGGCTCATCATGCTATGCTCTTAATGCTTAT 467
440 TTGTCCTCAATTCCTGACTTGTGTTCTCTGATGATGCTATGCTCTCCGAATGTTAT 499
468 GTCCTGCTGTGACAGCTTCGTAGATGATGATTTGGTATTTGGCTGTGCTGACGCTTGTG 527
500 GTGTGCTCTTAGAGAGGTGTATGATGATTTGGGGTGTGCTTGTGTACACTTGT 559
528 CCTTTGTACATTCAGAAATAGCCCCCTCGAGATTAGAGTTTGTCTGAATACACTACCA 587
560 CCTATTATATATCTGAGACTGCACCACTGAGATTAGGGGACTGTGAATACGCTACCG 619
588 CAATTCACTGATCAGAGAGAAATGTTCTGTCTACTGATGCTGTGTTGGATGCTCCCTG 647
620 CAGTTCACTGCTCTGAGGGAGTGTCTTATCTTACTGTATGTTTTCGAATGCTGTG 679
648 TCGCCATCACCCGATTGAGAAATATGCTGTGTGCTGCTCGGATACCTTCAATGTTCTTC 707
680 ATGCATCACCTAGCTGAGATTGATGCTGTGTCTTTCATCCCTTCTGCTTT 739
708 TTGTGTTGACAATATTTTATCTTCTGAACTCCAGATGCGCTGTTAGCAAGGTCGG 767


```
Db 740 TTCTCTCAGCGTCTTCTTCTGCCCCGAGTCCCCAAGGTGGCTCGTGAGCAAAAGTGA 799
QY 768 ATGGCAGAGGCAAAAAAGGTGTGCAAAAAGTTACG3GGGAAAGACGATGTCAGGTGA 827
Db 800 ATGCTTGAAGCAAAAGCGGTTCTTACAGACTCGTGGTCGGAAGATGTGTCTGGTAG 859
QY 828 TTGTCCCTTCTCTCGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGAGTACATC 887
Db 860 ATGGCTTTGTGTGAGGGTCTTGAATTGAGGTGAACAACCATAGAGAAATATATA 919
QY 888 ATTGAACCTGCCACCGAGGACGCCGATGATCTTGTACTGAACGGTGATAAGAACAAATC 947
Db 920 ATTGCTCCCGCGGATGAAGTACTGATGATCATGATATAGCTGTGGATAAGATCAAAAT 979
QY 948 ACACTTATAGGCGCTGAAGAGGCCAGTCATGGATTGCTGACCTTTAAAGGACCCATC 1007
Db 980 AAGTTATATGTTGAGAGAGAGGCTGAGTTGGTTGCTAGGCCAGTCAAAAGGA----- 1033
QY 1008 ATGCTTGAAGTGTGCTTCTCTGTCATCTCGTCANGGAGCATGTTGAACAGAG---T 1064
Db 1034 -----GGAAGCACTATGAGTGTGTTGTCTCGCCATGGAAGTACAATGAGCAGAGGCAA 1087
QY 1065 GTACCCCTTATGATCCGATTTGACACTTTTGTAGTTCATGAGAATATGCCCTCAA 1124
Db 1088 GGCTCATGTGATGATCCTCTGTTCACACTGTTGG3AGCGTTCAAGAGAGATGCCGAC 1147
QY 1125 GCTGAGGAAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCAGTGTACA 1184
Db 1148 ACT---GGAAGCATGAGGAGTGTCTTGTCCACATTTTGGAGATATGTTCAGTGTGGA 1204
QY 1185 GATCAGCATGCCAAAATAGACAGTGGGATGAAGAAATCTTCATAGGGATACGAGAG 1244
Db 1205 GGGAAATCAACCAAGACATGAAGATTGGGATGAAGA3AATCTTTGTGAGAAAGGTGAGAT 1264
QY 1245 TAGGCATCTGATGTGTCAGAGAGTGACTATGAGGACAACTTCCATAGCCCATTTGCTCC 1304
Db 1265 TATTCATCCGACCAT---GGAAGATTTCTGAAGATGATCTTCATCTCCGTGATCTCA 1321
QY 1305 AGGCAAGCAACAGGTGCGAAGGAAGACATTTGT3CACCAATGTTCAACCGTGAAAGTCT 1364
Db 1322 CGTCAAAACGACAGCATGGA---GAAGAATGCTCTCACACTGCTCATGAACTCTTCT 1378
QY 1365 TTGAGCATGAGAAGCAAAACCTCTTAGGGAGGGTGAAGTGTGTGAGCAGCATGAT 1424
Db 1379 ACCTTCAGACATGGAAGTCAAGTGCAGGGAGCTCAAGGGGAAGAGCGGTAGTATGGGG 1438
QY 1425 ATCGGTGGGGGATGGCAGCTTGTCTGAAATGTGTCAAGAAAGAGAGGTGAGAATGTACA 1484
Db 1439 ATTGAAGGTGATGGCAAGTGGCATGGAATGAC3GAAAGAGAGATGAATCGGGACAG 1498
QY 1485 AAGGAAGGTGTTTCAAAAGAGTCTACTTGCAACCAAGAGGAGTTCCTGGCTCAAGAAAG 1544
Db 1499 AAAGAAGGTGGGTTTAAACGGATATCTTGATCAAGAAAGTTTCCAGGATCTCGAGCT 1558
QY 1545 GGCCTCAATTTGTTCACTTCCCGGTGGTGGCGATGTTTGTGAGGGTAGTGAAGTTGTACAT 1604
Db 1559 GGCTCAATTTGTTCACTTGGTGGTGGTGAAGACCGGTGAGGCA--GATTTGTACAA 1615
QY 1605 GCTGCTGCTTTAGTAGTCAAGTCACTTTTCTCAAGGTTCTTGCTGAACCAAGCATG 1664
Db 1616 GCGTCTGCTTTGTTAGTCAACCAAGCATCTTTATTCAAAGACCTTCTCAAGAACATACA 1675
QY 1665 TCAGATGCTGCCATGTTCAACCAATCTGAGGTAGCTGCCAAAGGTTCAAGTTGAAGAT 1724
Db 1676 ATTGCTCTGCTATGTGATCAATCCGAACAACACT--AAAGGTCATTTGGCATGAT 1732
QY 1725 TTGTTTGAACCTGAGTGAAGGCGTGGCTGTAGTGGGTGTTGGAATTCAAGATCCTTCAA 1784
Db 1733 CTTCAATGATCCTGAGTCAAGCGTGCATTAAGTGTAGAGATTGGAAGTTCAAAATACTTCA 1792
QY 1785 CAGTTTGTGGAATAAACGGTGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGT 1844
```

```
Db 1793 CAGTTCTCAGGCATCAACGGAGTCTTTACTACACACCGCAATCTTGAGCAGCGGGT 1852
QY 1845 GTGGCAGTTATTTCTTCCAAATTTGGTCTCAGCTCCGCAATCAGCATCCATCTGTAGTAC 1904
Db 1853 GTCCGGATCCTACTATCGAACATGGGGATTAGTTCTTCTCAGCATCTTACTTATAAGT 1912
QY 1905 TCTCTCACTACCTTACTAATGCTTCTCCTTGCAATGGCTTTGCCATGCTGCTTATGATCTT 1964
Db 1913 GCATTGACAAACCTTTGTGATGTTACCTGCAATAGCTGTTGCAATGAGGCTCATGATCTT 1972
QY 1965 TCCGGAAGAGTTTGTGCTGCTAGGCACAAATTTCCAAATCTGATAGCATCTTACTATATC 2024
Db 1973 TCTGTCGAAGGACCTTGTCTTCTCACCACGATACCAATCCTGATAGCATCTTATGTT 2032
QY 2025 CTGTTGTCTCAATCTAATTGATTTGGGTACACTAGCCCATGCTTGTCTCCACCGTC 2084
Db 2033 TTAGTAATCTCAATCTTGTTCACATGAACAGCATGTGACAGCGGTCTTATCAACCGTA 2092
QY 2085 AGTGTATCTCTACTTCTGCTGCTTCTGTTAAGGATTTGTCCCATCCCAACATTTTA 2144
Db 2093 AGCGTTGTCTTACTTCTGCTTCTTCTGATGGATTTCGGTCTGCTCCAAACATCTC 2152
QY 2145 TGTGACAGATCTTTCCAAACCAAGGTTGCTGCTCTGTATTTGCCATTTGTGCTTTACA 2204
Db 2153 TGTTCAGAGATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCGCACTCAC 2212
QY 2205 TTCTGATCGAGATATCATGCTCACCTACAGCTTCTGATGCTGAATGCTATTTGA 2264
Db 2213 TTCTGATCTGTGACATAATGCTCACCTTACAGTCTCCCGTGTCTCAAAATCCATTGGA 2272
QY 2265 CTGCGGGTGTTCAGCATATATGACGTGATGCTTGAATTTCTTGTGTTGTTGCTTTC 2324
Db 2273 CTAGCTGTGTGTTGGAATGTACGCAATCGTATGTTCATTTCAATGGGTCTTGTGTT 2332
QY 2325 CTTAAGTCCCTGACAAAGGGGATGCCCTTGAGGTTATTACCGAATCTTTGACGTT 2384
Db 2333 ATTAAGTCCCGGAATAAGGCATGCCACTGAAGTCAATCAGAGATCTTTTCTGTT 2392
QY 2385 GGTCCGAAGCAAGCGCTGC 2404
Db 2393 GGAGCTAGACAAAGCTGAAGC 2412
```

```
RESULT 7
AX506620
LOCUS AX506620 2190 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1315 from Patent WO0216655.
ACCESSION AX506620
VERSION AX506620.1 GI:23387857
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 1315 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source
1. .2190
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 529 a 439 c 572 g 650 t
ORIGIN
Query Match 31.5%; Score 875.8; DB 6; Length 2190;
Best Local Similarity 64.9%; Pred. No. 2,7e-219;
Matches 1449; Conservative 0; Mismatches 722; Indels 60; Gaps 8;
```

QY	177	ATGTCGGGGGCTGTTCTTGTCGCCCATAGTCGCCCTCCATCGGCAATCTATTGCAGGGGTG	236
Db	1	ATGAGTGGAGCTGTGCTTGTTGCTATTGCTGCTGCTGTGGCAACTTGTACAGAGATGG	60
QY	237	GACAATGCCACCATCGCAGCTGCTGTTCTGTATATAAGAGAAATTTCAATTGCCAAAT	296
Db	61	GATAACGCACTATTGCGAGAGCTGTGTGTACATAAAAAAGAGTTAATTGGAGAGT	120
QY	297	GAGCCCACTGTGAGGGACATAATTGTCAATGTCACTTAATCGCGCCACCACCATCGTTACT	356
Db	121	AATCCATCAGTGAAGGTCTAATTGTGCGCATGTCACTTAATTGTGCTACTCTGTATTACA	180
QY	357	ACATTTCCGGGCCATTATACAGCTCGATTGGCCGACGCCCTATGCTTATTCTCTTCA	416
Db	181	ACATGCTCTGAGGGGTAGCTGATGTGGCTTGTGCGCCGTCATGCTAATATTGTCTCA	240
QY	417	ATTCTGTACTTCTTACGCGGCTCATCATGCTATAGTGTCTCTAATGTCTATGTCTGCTG	476
Db	241	ATTCTTACTTGTGTGTTCTCTAGTAATGCTATAGTGTCTCGAATGTTATGTGTGCTC	300
QY	477	TTGGCACGCTTGTAGATGGAATTGGTATTGGCTTGGCTGTCAAGCTTGTGCTTTGTAC	536
Db	301	TTAGGAAGGTGTATGATGGAATTGGGGTTGGTCTTGTGTGTCACTTGTCTCTATTAT	360
QY	537	ATTTCAGAATAAGCCCTTCGGAGATTAGAGTTTGTCTGAATACTACCACAATTCACT	596
Db	361	ATATCTGAGACTGCACCACCTGAGATTAGGGGAACTGTTGAATACGCTAACCGCAGTTCACT	420
QY	597	GGATCAGGAGGAATGTTCTTGTCATACCTCATGTGTGTTGGATGTCCCTGTGCGCATCA	656
Db	421	GGCTCTGAGGAGATGTTCTTAATTACTGTATGTTTCGAAATGTCGTTGATGCCATCA	480
QY	657	CCCGATTGAGAAATTATGCTTGTGCTGCTCGGATACCTTCATTGTTCTTGTGTTG	716
Db	481	CCTAGCTGAGATGATGCTGTGTGCTCTTTTTCATCCCTTCCCTTGTCTTCTTCCTC	540
QY	717	ACAATATTTTATCTTCCGTAATCTCCAAGATGGCTCGTTAGCAAAAGCTCGGATGGCAGAG	776
Db	541	ACGGTCTTCTTCTTGCCCGAGTCCCAAGGTGGCTCGTGACCAAAAGCTCGAATGCTTGAA	600
QY	777	GCAAAAAAGGTGTTGCAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATTGTCCCTT	836
Db	601	GCAAAAGCGGTTCTTCAGAGACTGCTGTGCGCAAGATGTCTGTGTAGATGGCTTTG	660
QY	837	CTTCTCGAAGGTTGAGGTTGAGAGAGACACTTCCATTGAAGATACATCATTTGACCT	896
Db	661	TTGTTGAGGTTCTTGAAATTGAGGTGAACAACCATAGAGAAATATATATTGTGCC	720
QY	897	GCCACCGAGCAGCCGATGATCTTGTACTGACGGTATAGGAACAATCACACTTAT	956
Db	721	GCGGATGAAGTTACTGATGATCATGATATAGCTGTGATTAAGATCAATTAAGTTATAT	780
QY	957	GGGCCTGAAGAGGCCAGTCATGGAATTGCTCGACCTTCTTAAGGGAACCCATCATGCTTGA	1016
Db	781	GGTGCAGAAAGGGCTGAGTTGGGTTGCTAGGCCAGTCAAAAGG-----GGA	828
QY	1017	AGTGTCTTCTCTTGCACTCTGTCATGGGAGCATGTGTGAACCAAG--TGTACCCCTT	1073
Db	829	AGCACTATGAGTGTGTTGTCTCGCCATGGAAGTACAATGAGCAGAGGCAAGGCTCATTG	888
QY	1074	ATGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAAATATGCTCAAGCTGAGGA	1133
Db	889	ATTGATCCTCTTGTCACTGTTGGGAGCGTTCACGAGAGATGCCGACACT--GGA	945
QY	1134	AGTATGAGGAGCACATTGTTCCAACCTTTGGAAGTATGTTCAGTGTACAGATCAGCAT	1193
Db	946	AGCATGAGGAGTGCCCTTGTCCACATTTTGGAGTATGTTCAGTGTTGAGGGGAATCAA	1005
QY	1194	GCCAAAATGAGCAGTGGGATGAAGAATCTTCATAGGGAATGACGAGAGTACGCATCT	1253
Db	1006	CCAAGACATGAAGATTGGGATGAAGAATCTTGTGAGAGAGGTGAGGATATCATCC	1065

QY	1254	GATGGTGCAGAGGTGACTATGAGACAATCTCCATAGCCCATGTGCTGCCAGGCA	1313
Db	1066	GACCAT--GGAGATGATTTGAAAGATGATCTTCATCTCCGTTGATCTCAGTCAACG	1122
QY	1314	ACAGTGCAGAGGAAGACATTGTGCACCATGCTCACCGTGGAAGTGTGAGCATG	1373
Db	1123	ACAAGCATGGA--GAAAGACATGCCCTCACACTGCTCATGGAACCTTTTCACTTCAGA	1179
QY	1374	AGAAGCAAAACCTCTTAGGGGAGGTGAGATGTTGTGAGCAGCAGCATGATATCGGTGG	1433
Db	1180	CATGGAAGTCAAGTGCAGAGGAGCTCAAGGGGAGAGACGGGTAGTATGGGATTGAGGT	1239
QY	1434	GGATGGCAGCTTCTTGGAAATGTCAGAGAAGAAAGTGAGATGTTAGAAAGAGT	1493
Db	1240	GGATGGCAAGTGGCATGGAATGGAACGGAAAGAGAAGATGAATCCGGACAGAAAGAA	1299
QY	1494	GGTTTCAAAAGAGTCTACTTGCACCCAAGAGGAGTTCCTGCTCAAGAGGGGCTCAAT	1553
Db	1300	GGTTTC-----CCAGGATCTCGACGTGGCTCAATT	1329
QY	1554	GTTTCACTTCCCGGTGCTGCGCATGTTTGTGAGGTAAGTATGATTTGTACATGCTGCT	1613
Db	1330	GTTTCATTGCTGTGTGTGATGGAACCGGTGAGCA--GATTTGTACAAAGCTTGTCT	1386
QY	1614	TTAGTAAGTCAATGACACTTTTCTCAAAGGCTTGTCTGAACCAAGCATGTAGATGCT	1673
Db	1387	TTGGTTAGCCAAACAGCTCTTTATTCCAAAGACCTTCTCAAAGACATACAAATTTGCTCT	1446
QY	1674	GCCATGTTTCAACCATCTGAGGTAGCTGCCAAAGTTTCAAGCTTGGAAGATTTGTTGAA	1733
Db	1447	GCTATGTACATCCATCCGAACAACACT--AAAGGTCAATTGTGGCATGATCTTCATGAT	1503
QY	1734	CCTGAGTGAAGCGCTGCCCTGTAGTCCGTTGTTGAAATTCAGATCCTTCAACAGTTTGCT	1793
Db	1504	CCTGGAGTCAAGCGTGCAATTAGTCGTAGGAGTTGSACTTCAATACTTCAGCAGTTCTCA	1563
QY	1794	GGAAATAAACGGTGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGTGTGCAAGTT	1853
Db	1564	GGCATCAACGAGTCTTTACTACACACCGCAAAATCTTGAGCAGCGGCTGTGGGATC	1623
QY	1854	ATTCTTCCAAATTTGTTCTCAGCTCGGCATGCATCCATCTTGATCAGTTCTCTCACT	1913
Db	1624	CTACTATGAAACATGGGGAATTAGTCTTCCCTCAGCATCTTACTTATAAGTCAATTGACA	1683
QY	1914	ACCTTACTAATGCTTCTTGCAATTGGCTTTGCCATGCTGCTTATGATCTTTCCGAAGA	1973
Db	1684	ACCTTTGTGATGTTACTGCAATAGCTGTGCAATGAGGCTCATGGATCTTTCTGTGCA	1743
QY	1974	AGGTTTGTGCTGCTAGGCACAATTCCAAATCTTGATAGCATCTCTAGTTATCTGTGTGTG	2033
Db	1744	AGGACCTTGCTTCTCAACCAAGATACCAATCTGATAGCATCTTATGTGTTTAGTAATC	1803
QY	2034	TCCAATCTAATTGATTTGGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAAGTGTATC	2093
Db	1804	TCAATCTTGTTCATGACGAAACAGCATTTGTGACGCGGTCTTATCAACCGTAAGCGTTGTG	1863
QY	2094	GTTTACTTCTGCTGCTTGTATGGGATTTGTTGCCATCCCAACATTTTAATGTGACAG	2153
Db	1864	CTTACTTCTGCTTCTTGTGATGGGTTTGGTCTGCTCCAAACATCCTGTGTCAAG	1923
QY	2154	ATCTTTCAAACAGGGTTCGTGGCTCTGTATATGCCATTTGTGCTTACATCTGGATC	2213
Db	1924	ATTTTTCCAACTCGAGTCCGGGAATCTGCATGSCCATCTGCGCACTCACCTTGTGATC	1983
QY	2214	GGAGATATCATCTGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTTGACTGGCGGT	2273
Db	1984	TGTGACATTAATCTGCTTACAGTCTCCCGGTGCTCAAAATCCATTGGACTAGCTGGT	2043
QY	2274	GTTTTCAGCATATATGACAGTGTGATGCTTGAATTTCTTTGTGTGCTTCTTAAAGGTC	2333
Db	2044	GTGTTTGAATGTACGCAATCGTATGTTGCAATTTCAATGGGCTTTGTGTTCATTAAAGTC	2103
QY	2334	CCTGAGACAAAGGGGATGCCCTTGAGGTTATTAACGAATTTCTTTGACAGTTGCTGCGAAG	2393

Db	2104	CCGAAACTAAGCATGCCACTTGAAGTCATCACAGAGTCTTTCTGTGAGCTAGA	2163
QY	2394	CAAGCGGCTGC	2404
Db	2164	CAAGCTGAAGC	2174
RESULT 8			
ATH532570		2190 bp	mRNA linear
LOCUS			PLN 11-DEC-2002
DEFINITION		Arabidopsis thaliana mRNA for monosaccharide sensing protein 2	
ACCESSION		AF532570	
VERSION		Ad532570.1	GI:26800694
KEYWORDS		monosaccharide sensing protein 2; mssp2 gene.	
SOURCE		Arabidopsis thaliana (chale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	1	Stramie,C., Tjaden,J., Trentmann,O., Emmerlich,V. and Neuhaus,E.	
TITLE		A novel family of plant monosaccharide transporters is involved in	
JOURNAL		a new type of eukaryotic sugar sensing	
REFERENCE	2	Neuhaus,E.	
AUTHORS		Neuhaus,E.	
TITLE		Direct Submision	
JOURNAL		Submitted (04-DEC-2002) Neuhaus E., Plant Physiology, University of	
REFERENCE		Kaiserslautern, Erwin-Schroedinger-Str. 22, 67663 Kaiserslautern, GERMANY	
FEATURES			
source		Location/Qualifiers	
	1..2190	/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
gene		/db_xref="taxon:3702"	
	1..2190	/gene="mssp2"	
CDS		1..2190	
		/gene="mssp2"	
		/codon_start=1	
		/product="monosaccharide sensing protein 2"	
		/protein_id="CAD58692.1"	
		/db_xref="GI:26800695"	
		/translation="MSGAVLVAIAAVGNLQGMNATIAGAVLYIKKEENLESNPSV	
		EGLIVAMSLIGATLITTCGGVADWLGRRMILSSILYFVGLVLMSPENVYVLLG	
		RLUDFGVGLVTVLVEIYISETAPPEIRKILNTLPQFTSGGMFLSYCMVFGMSLMS	
		PSWRLMLGLVLPISLVFPLTVFPLPESPRMLVSKGRMLEAKRVLQRLRGEDVSGEM	
		ALLVEGLIGGETTIEEYIIGPADEVTDIHDAVDKQIKLYGAEGLSNVAPVKGG	
		STMSVLRSRHSFMSRQSLIDPLVTLFGSVHEKMPDTGMSRALFPHFGSMFSVGN	
		QPRHEDWDEENLVGEEDYPSDHGDSEIDLHSPILSRQFTSMKMDPHTAGTLSTF	
		RHGSQVGAQGEAGSMGIGGQWAVAMKTEREDESGQKEEGFPGSRGSIVSLPGGD	
		GTGEADFVQASALVQPALYSKDLKEHTIGPAMVHPSETTKGSIWHDLDHFGVKRAL	
		VVGVLQIILQFSGINGVLYYTPQILEQAGVGILSNMGISSSASLISALTTFVML	
		PAIVAMRLMDLSGRRLTLTTIPILIASLVLVISNLVHMNSIVHAVLSTVSVLYF	
		CFVMGFGAPNLICSEIFPTRVRGICIAICALTFWICDIIIVTYSPLVLKSIGLAGV	
		FGMYAIVCCISWVFETFKVPETKGMPLVITEFESVGARQAEAAKNE"	
BASE COUNT	529 a	439 c	572 g 650 t
ORIGIN			
Query Match	31.5%;	Score 875.8;	DB 8; Length 2190;
Best Local Similarity	64.9%;	Pred. No. 2.7e-219;	
Matches 1449;	Conservative 0;	Mismatches 722;	Indels 60; Gaps 8;
QY	177	ATGTCCGGGGCTGTCTTGTCCGCAATAGTCGCCCTCCATCGGCAATCTATTGCAGGGGTGG	236
Db	1	ATGAGTGGAGCTGTGCTTGTGCTATTGCTGCTGCTGCTTGCGCACTTGTTACAAGATGG	60
QY	237	GAGCATGCCACCATCGCAGCTGCTTCTGTATATAAAGAAATTTCAATTGCAGAAAT	296
Db	61	GATTAACGCAACTATTGCAGAGCTGTGTGTACATMAAAAGAGTTTAATTGGAGAGT	120

QY	297	GAGCCACTGTGAGGGACTAATGTGTCAATGTCACCTTATCGGCCACCATCGTACT	356
Db	121	AATCCATCAGTGGAGAGTCTAATGTGTGCGGATGTCACCTATTGTGTGCTCTGATTACA	180
QY	357	ACATTCTCCGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTTCA	416
Db	181	ACATGCTCTGAGGGGTAGCTGATTGGCTTGGTCGCCGTCCCATGCTAATATGTCTCTCA	240
QY	417	ATTCTGTACTTCTTCAGCGGCTCATCATGCTATGCTCTCTAATGTCATGTCCTGCTG	476
Db	241	ATTCTCTACTTGTGTGTTCTCTAGTAATGCTATGCTCTCCGAATGTTATGTTGCTC	300
QY	477	TTCGACGCTTCTGATGATGATTGCTATGCTTGGCTGTCCACGCTTGTGCTTGTAC	536
Db	301	TTAGGAAGTGTGTAGATGATTTGGGTTGGTCTGTGTGTCACACTTGTCTATTAT	360
QY	537	ATTTCAGAAATAGCCCTTCGAGATTAGAGTTTGCTGAATACACTACCAATTCAGT	596
Db	361	ATATCTGAGACTGCACCACTGAGATTAGGGGACTGTGAATACGCTACCGAGTTCAC	420
QY	597	GGATCAGAGGAATGTTCTGTCTCATCTGCATGCTGTTGGAGTGTCCGTGCCCATCA	656
Db	421	GGCTCTGAGGGAGTCTTCTATCTTACTGTATGCTTTTCGGAATGTCTGATGCCATCA	480
QY	657	CCCGATTGAGAAATATGCTTGTGTGCTGCTCGGATACCTTCATGTTCTTCTTGTG	716
Db	481	CCTAGCTGAGATTGATGCTTGTGTGCTTTCATTCCTTCCTTCTTCTTCTCTC	540
QY	717	ACAATATTTTATCTCTCTGAATCTCCAGATGCTGTTAGCAAAAGTGGATGGCAGAG	776
Db	541	ACGGTCTTCTTCTTCCCGAGTCCCAAGTGGCTGTGAGCAAAAGTGAATGCTTGAA	600
QY	777	GCAAAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGCTCAGGTGAATGTCCCT	836
Db	601	GCAAAAGCGGTTCTCAGAGACTGCGTGTGCGGAAGATGTGTGTGAGATGCTTGTG	660
QY	837	CTTCTCGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGATACATGATGGACCT	896
Db	661	TTGTTGAGGCTTGGAAATGAGAGTGAACACACATAGAGAAATATATGATGCTCC	720
QY	897	GCCACCGAGGCGGACATGATCTTGTATCTGACGCTGATTAAGAACAAATCACACTTAT	956
Db	721	GCGGATGAAGTTACTGATGATCATGATATAGCTGTGGAATTAAGATCAAAATTAATAT	780
QY	957	GCGCTGAAGAGGCGGACATGATGATGCTGCACTTCTTAAGGAGCCCATGCTTGA	1016
Db	781	GCTGAGAAAGAGGCTGAGTTGGTGTGCTAGGCCAGTCAAAAGA-----GGA	828
QY	1017	AGTGTGCTTCTCTTCATCTGTCATGAGGAGCATGTGAACAGAG--TGTAACCTT	1073
Db	829	AGCACTATGAGTGTGTTGCTCTCGCATGAAATGAATAGACAGAGGCAAGGCTCATTTG	888
QY	1074	ATGATCCGATTGTGACACTTTTGTAGTGTCCATGAGATATGCTCAAGCTGAGGA	1133
Db	889	ATTGATCTCTTGTGACACTGTTGGAGCGTTCAGAGAGATGCCGACACT--GGA	945
QY	1134	AGTATGAGAGCACATTTGTTCCAACTTTGGAAGTATGTTCAGTGTACAGATCAGCAT	1193
Db	946	AGCATGAGAGTGCCTTGTTCACCATTTTGGAGTATGTTCAGTGTGAGGGAATCAA	1005
QY	1194	GCCAAAATGAGCAGTGGGATGAAGAGATCTTCATAGGGATGACGAGAGTACCATCT	1253
Db	1006	CCAAGACATGAAGATTGGGATGAAGAGATCTTGTGGAGAGGTTAGGATTAATCCATCC	1065
QY	1254	GATGTGAGAGGATGACTATGAGACAAATCTCATAGCCCATGTGCTGTCCAGGAGCA	1313
Db	1066	GACCAT--GGAGATGATTTGAAAGATGATCTTCATCTCCGTGATCTCAGTCAAAAG	1122
QY	1314	ACAGTGCAGAGGAGAGACATTTGTGCACCAATGCTCACCGTGAAGTCTTGAAGCATG	1373
Db	1123	ACAAGCATGGA--GAAGAATGCTCTCACACTGCTCATGGAATCTTTCTTACTTCA	1179
QY	1374	AGAAGCAAAACCTTTTAGGGGAGGTTGAGATGCTGTGAGCAGCACTGATATCGGTGG	1433

Db	1180	CATGGAAGTCAAGTGCAGGAGCTCAAGGGAGGAGCGGGTAGTATGGGATTTGAGGT	1239
QY	1434	GGATGGCAGCTTGCTTGAATAATGCTCAGAGAAGGAGTGAATGTAGAAAGAGGT	1493
Db	1240	GGATGGCAAGTGGCATGGAATAAGACGGAAAGAGAATGAATCCGGACAGAAAGAGAA	1299
QY	1494	GGTTCAAAAAGTCTACTTGCAACCAAGAGGAGTCTCGGCTCAAGAGGGGCTCAATT	1553
Db	1300	GGTTTC-----CCAGGATCTCGACGTGGCTCAATT	1329
QY	1554	GTTTCACTTCCCGGTGGCGGATGTTTGGAGGGTAGTGTGTATCATGCTGCTGCT	1613
Db	1330	GTTTCATTGCTGTGTGATGGAACCGGTGAGGCA---GATTTGTACAAGCGTCTGCT	1386
QY	1614	TTAGTAAGTCAAGTCAAGTCTTCTCAAAGGCTTGTGTAACCAAGCATGTCAAGTCT	1673
Db	1387	TTGGTTAGCCAACCAAGCTCTTATTCCAAAGACCTTCTCAAAGAATACATATTGCTCT	1446
QY	1674	GCCATGCTTCAACCATCTGAGGTAAGCTGCCAAAGTTCAAGTTGGAAGATTTGTTGAA	1733
Db	1447	GCTATGTTACATCCATCCGAACAACCT--AAAGGTCATATTGGCATGATCTTCATGAT	1503
QY	1734	CCTGAGTGAAGCGGCTGCTGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1793
Db	1504	CCTGAGTCAAGCGTGCATTAAGTGTGAGAGTGTGACTTCAATTAATCTTCAAGCAGTCTCA	1563
QY	1794	GGAATAAACGGTGTCTGTACTATAACCCACAATCTTGAGCAAGCTGTGTGCGAGTT	1853
Db	1564	GGCATCAACGAGGTTCTTACTACACACCGCAATCCTTGAGCAGCGGCTGTGCGATC	1623
QY	1854	ATCTTTTCCAATTTGGTCTGACTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACT	1913
Db	1624	CTACTATCGAATGAGGATGATGTTCTTCTCAGCATCTTACTATATAGTCATGACA	1683
QY	1914	ACCTTACTAATGCTTCTGCAATGGCTTTGCCATGCTGCTTATGATCTTCCGGAAGA	1973
Db	1684	ACCTTTGTGATGTTACCTGCATAAGCTGTTGCAATGAGGCTCATGATCTTCTGTGCA	1743
QY	1974	AGGTTTGTGCTGCTAGGACAAATTTCCAATCTTGATAGCATCTTCTAGTATCTCTGTTG	2033
Db	1744	AGGACCTTGCTTCTCACCACGATACCAATCTGATAGCATCTTCTATTGTTTATGATATC	1803
QY	2034	TCCAATCTAATGATTTGGGTAACAAGCCCATGCTTGTCTCCACCGTCAGTGTATC	2093
Db	1804	TCAATCTTGTTCATGATGAACAGATTTGTGCACGCGCTTATCAACCGTAAGCGTTGTG	1863
QY	2094	GTTCTACTTCTGCTGCTTCTGTAAGGATTTGGTCCCATCCCAACATTTTATGTGAGAG	2153
Db	1864	CTCTACTTCTGCTTCTTCTGTAAGGTTTGGTCTCTGCTCCAAACATCTCTGTTCAGAG	1923
QY	2154	ATCTTTCCAACCAAGGTTCTGCTGCTCTGTATTTGCCATTTTGTGCTTTACATTTGATC	2213
Db	1924	ATTTTTCCAACTGAGTCCGCGGAATCTGCATCGCCATCTGCGCACTCACCTTCTGATC	1983
QY	2214	GGAGATATCATGCTCAGCTTACAGCTTCTCTGTGATGCTGAATGCTATTGGAGTGGCGGT	2273
Db	1984	TGTGACATTAATCTGCTACCTTACAGTCTCCCGTGTCTCAATTCATTTGACTAGCTGCT	2043
QY	2274	GTTTTCAGCATATATGAGTGCATGCTGATTTCTTGTGTTGTTGTTGTTGTTGTTGTTG	2333
Db	2044	GTTTTCAGCATATATGAGTGCATGCTGATTTCTTGTGTTGTTGTTGTTGTTGTTGTTG	2103
QY	2334	CCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATCTTTGCAAGTGTGCGAAG	2393
Db	2104	CCGGAACATAAAGGCATGCCATTAAGTATCATCAGAGTCTTCTTCTGTGAGCTAGA	2163
QY	2394	CAAGCGCTGC 2404	
Db	2164	CAAGCTGAAGC 2174	

RESULT 9

QY	415	CAATCTGTACTTCTTCAAGCGGCTCATCATGCTATGCTCTCTAATGCTTATGCTCTGC	474
Db	264	CGAACTGCTCTTCTTGGCTCCAAACGATGCTCTGCGCGCCCAACGCTTATGTTGCTC	323
QY	475	TGTTGGACGCTTCTGATGATGATTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT	534
Db	324	TCTTGGCCAGGCTCATGACGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	383
QY	535	ACATTTCAAGATAGCCCTTCCGAGATTAGGTTGCTGATACACTACACAATTCA	594

FEATURES

source

1. 2516

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/culivar="barke"

/db_xref="taxon:112509"

/tissue_type="developing caryopsis"

/clone_lib="lambda ZAP2"

/dev_stage="0-15 days after flowering"

1. 2516

/gene="stp2"

77. 2338

/gene="stp2"

/codon_start=1

/product="sugar transporter"

/protein_id="CAD58959.1"

/db_xref="GI:26986188"

/translation="MRGGWRSWRCGRWRARRSSRLRGWRRRRAGAGARKDLAA

LQAEWKSLEELVMAAAKLVFLANVLMAPNVVLLARLDGFIGLATVPL

YISETAPTDI RGLNTLPFGSGGFLSYCMVFTMSLMPDPRIMLGVLISPLMY

FALTVPYIPESPRMLVSKGRMAEAKVLRGRREDVSGEMALVEGLGVKTHIEY

IIGPDELADDDGLADDEKLVGAEBGVSWIARPVGGGSGALGSLMRHGSMT

SQGSLLVDPLVTLFGSVHEKMPVWMSRSTLFPNFGSMFSVAEQOAKADWDAESH

DDEDYASDHGADDIEDNLNPLISRQATSVEKEIAAPHGSIIMGVSSSMOGDAVS

SMGIGGWQLAWKTEREGADGRKEGAQRIYLHEEVSQDRGSILMPGGDIPGG

EYIQAAALVSPALYSKDLIEQOLAPAMVHPSEAVAKGTMAELFEPGVKHALFVGI

GLQILOEFAGINGVLVYTPQILEQAGVILSNIGLSSSASILISALTLLMLGYIS

DRIGARAITAASFMYCSIPALFFYRRAILLVNVLDVGTMTAAISTISIVYFCFV

MGPIPEGSTSSVKRIDTDRAANAASLAKALLTFWIGDIIVTYTLPMNLNIGLAGVFIY

AVVCMIAFVFMKVPETKGMPLVITTFEFSVGAKQKEATDEVAEKIAESRPSSECR

SPVDYVQDHVMEV"

BASE COUNT

567 a 524 c 774 g 651 t

ORIGIN

Query Match 29.0%; Score 806.2; DB 8; Length 2516;

Best Local Similarity 65.1%; Pred. No. 6.1e-201;

Matches 1294; Conservative 0; Mismatches 668; Indels 27; Gaps 6;

Db 384 ACATTTCCGAGACCCCGCTGACATTAAGAGGCTGCTGAACACGCTGCCGACATTCA 443
QY 595 GTGGATCAGGAGGAATGTTCTTGATACATGATGTTGGGATGTCCTGTGCGCAT 654
Db 444 GTGGGTCAAGAGGATGTTCTTCTACTGATGTTGTTACCATGTCTCCATGCGCC 503
QY 655 CACCCGATTGGAGAATTATGCTTGCTGCTCGGATACCTCATTTGTTCTTTGGTT 714
Db 504 AGCTGACTGAGAAATCATGCTGGGGGTTTGTGATCCCGTCGCTTATGATTTTGCA 563
QY 715 TGACATATTTTATCTTCCGAACTCCAGATGCTCGTTAGCAAAAGTCGATGCGAG 774
Db 564 TGACTGCTTTTATTTGCCCCGAATCTCCAGATGCTTGAGCAAGGAAGATGGCTG 623
QY 775 AGGCAAAAAAGGTGTGCAAAAGTTACGGGGGAAAACGATGTTCTAGGTGAATTGTC 834
Db 624 AGGCCAAGCGATATTGCAAAAGACTGCGGGGAAAGGAAAGATGTTCTAGAGAGATGCCC 683
QY 835 TTCTTCTGAAGGGTTGAGTTGAGAGACACTTCCATTGAAGATCATCATGAC 894
Db 684 TTCTTGTGAAGGATTGGGTGT---GAAAAGACACACATTGAGGAATACATAATTGAC 740
QY 895 CTGCCACCGAGCGACCGCATGATCTTGTACTGACGATGATAAGAAACAATCACACTT 954
Db 741 CTGATGATGAGCTGCTGATGACGGGCTGCTCCA---GATCAAGAGAAGTTGAACCTCT 797
QY 955 ATGGCCTGAAGAAGGCCAGTACATGATGCTGCACTTCTAAG-----GACCCATCA 1008
Db 798 ATGAGCTGAAGAAGGGGATCTTGATCGCCCCGTCTGTAGGGGTGCGGCCAAGTG 857
QY 1009 TCGTTGGAAGTGTCTTCTCTGCACTCTGTCATGAGAGCATGTTGAACCAAGATGAC 1068
Db 858 CACTTGGAAGCGCTTGGGCTCATGCTCTGTCATGAGATGTTAGTCAAGGTAAAT 917
QY 1069 CCCTTATGATCCGATTGTGACATTTTGGTAGTGTCCATGAGAATATGCTCAAGCTG 1128
Db 918 CTCTGTGACCCCTGTGACTCTCTTCCGAAGTGTCCATGAGAAGATGCCGTGAAT 977
QY 1129 GAGGAAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCAAGTACAGATC 1188
Db 978 TGGGAGCATGCGAGACCTTGTCTTAATTTGAGCATGTTAGTGTGCTGAC 1037
QY 1189 AGCATGCCAAAAATGACAGTGGATGAAGAAATCTTCATAGGATGACGAGAGTACG 1248
Db 1038 AGGACAGGCTAAAGCTGACTGGATGCTGAGA---GTCAAGGATGATGAAGATATG 1094
QY 1249 CATCTGATGTCAGAGAGTGAATAGAGACAATCTCCATGCCCCATTGCTGTCCAGC 1308
Db 1095 CATCGATCATGCTGCTGATGACATTAAGGATTAACCTCAATAGCCCCCTTATTTCTG 1154
QY 1309 AGGCAACAGGTGCGGAAGGGAAGACATTTGTGACCATGTTCAACCGTGGAAGTCTTGA 1368
Db 1155 AAGCGACAAGCGTGAGGGTAAAGATTTGCTGCACCTCATGGAAGCATTAATGGGT--- 1210
QY 1369 GCATGAAAGGCAAAACCTTTAGGGGAGGGGTGAGATGTTGAGCAGCATGATATCG 1428
Db 1211 ---GCTGTGAAAAGTAGACATGACAGGAGGAGACGACATTAAGCAGCATGGCATTG 1265
QY 1429 GTGGGGATGCGAGCTTGTGAAATGCTCAGAGAAAGAGAGTGAAGTGTAGAAAG 1488
Db 1266 GTGGGGGTGGCAGTTAGCTTGAAGTGAAGTGAAGAGAGAGTGAAGTGGCGAAAG 1325
QY 1489 AAGTGTGTTCAAAAGAGTCTACTTGCACCAAGAGAGAGTTCCTGGCTCAAGAGGGCT 1548
Db 1326 AAGGTGGCGCCAGCGTATTTACTTGCATGAGAGAGAGTGTCAAGTATCGAGAGGCT 1385
QY 1549 CAATTGTTTCACTTCCCGGTGTGCGCATGTTTGAAGGATGAGTTGTATCATGCTG 1608
Db 1386 CTATATTGTTATGCGAGAGGT---GATATTCTCTGCTGTGATATATCCAGGAG 1442
QY 1609 CTGCTTAAAGTCAAGTCAAGCATTTTCTCAAAAGGCTCTTGTCTGAACCAAGCATGTCAG 1668

Db 1443 CCGCTCTGTGAGCCAACTGCTCTTACTGAAAGACCTGATAGACAGACCTTGCTG 1502
QY 1669 ATGCTGCCATGTTCAACCATCTGAGGTAGCTGCCAAAGTTACGTTGAAGAATTTGT 1728
Db 1503 GTCCAGCCATGTTATTCATCCGAGGAGTGGCCAAAGGTTACAAAGTGGGCAACTAT 1562
QY 1729 TTGAACCTGAGTGAAGGCGTCCCTGTTAGTCCGTGTGGAATTGAGATCTTCAACAGT 1788
Db 1563 TTGAACCTGAGTGAAGCATGCACTGTTGTTGGCATTTGATTACAGATCTTGCAAGAGT 1622
QY 1789 TTGCTGAATAAACGGGTCTCTGTACTATACCCCAAAATCTTGAGCAAGCTGTTGG 1848
Db 1623 TTGCGGTATCAACGAGTCTGTATTACACGCTCAGATACTTGAGCAAGAGTGTG 1682
QY 1849 CAGTTATCTTCCAAATTGGTCTCAGCTCGGATCAGCATCCATCTTGATCAGTTCTC 1908
Db 1683 GGATTTCTTATCAAAATGAGCTAAGCTCTTCTGATCTATCTTATAGTGCTT 1742
QY 1909 TCACTACCTTACTAATGCTTCTGATTTGCCATGCTGCTTATGATCTTTCCG 1968
Db 1743 TGACAACCTTGCTGATGCTTGGTTACATTTCAACCGAATTTGTGCAAGAGCTATACG 1802
QY 1969 GAAGAAGTTTGTGCTGCTAGGACAATTCATCTTGATAGCATCTTATCTCTG 2028
Db 1803 CTGCGAGTTTCACTGATGCTGCTCAATTCAGCTTGTCTTCTACCGTAGAGCTATCTTG 1862
QY 2029 TTGTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCAGCTCAGTG 2088
Db 1863 TTTTGTGAATGTTCTGATGTGGAAACCATGTGCAAGCCGCTCTCTCAACGATCAGTG 1922
QY 2089 TTATGCTACTTCTGCTGCTGCTGTTATGGATTTGCTCCATCCCAACATTTATGTG 2148
Db 1923 TCATAGTCTACTTCTGCTTCTGTCATGGGTCCAAATCCGGAAGATCCCAAGCTCGG 1982
QY 2149 CAGAGATCTTCCAAACAGGGTTCGTGCTGCTGTAATGCCATTTGTGCTTATCATTTCT 2208
Db 1983 TCAAGATTGATACAGATAGAGCGAAGCAGCATGCTCGCTAAAGCATTAACCTTCT 2042
QY 2209 GGATCGAGATATCATGCTCACCTACAGCTTCTGTGATGCTGAATGCTATGGAAGTGG 2268
Db 2043 GGATTTGTGACATCATGCTGACATACACTCTTCCGTGATGCTCAACGCCATGGAAGTGG 2102
QY 2269 CCGGTGTTTCCAGCATATATGCAAGTGTGATGCTGATTTCTTGTGTTGCTTCTCTTA 2328
Db 2103 CTGGAATTTCCGAATATATGCTGCTGTTGCAATGATGCTTGTGATTTGCTTACATGA 2162
QY 2329 AGGTCCCTGAGACAAAGGGATGCCCTTGAGTTATTAACGAATCTTGGCAGTTGGTG 2388
Db 2163 AGGTCCCTGAGACAAAGGGCATGCCCTGAGGTATCACCGAGTTCTTCTGCTGCGGG 2222
QY 2389 CGAAGCAAG 2397
Db 2223 CGAAGCAGG 2231

RESULT 10
AX412656
LOCUS AX412656 2205 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 420 from Patent WO0222675.
ACCESSION AX412656
VERSION AX412656.1 GI:21445114
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
AUTHORS Plant genes, the expression of which are altered by pathogen
TITLE infection
JOURNAL Patent: WO 0222675-A 420 21-MAR-2002;
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT

CHAPBL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (US) ; Eulgem, Thomas (US)
Location/Qualifiers
1.2205
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 521 a 524 c 521 g 639 t
ORIGIN

Query Match 25.2%; Score 700; DB 6; Length 2205;
Best Local Similarity 61.2%; Pred. No. 6e-173;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;

```
OY 177 ATGCGGGGGCTGTTCTGTGCGCAATAGTCGCCCTCCATCGGCAATCTATTGCAGGGGTGG 236
Db 1 ATGAAGGAGCGACTCTGTTGCTCTCGCCGACAAATCGCAATTTCTTACAAAGATGG 60
OY 237 GACAAATGCCACCATCGCAGCTGCTGTCTGTATATTAAGAAGAAATTTCAATTGCAGAAAT 296
Db 61 GACAAATGCCACCATGCTGAGCTATGCTTTATATCAACAAGAACTTGAATCTAACCA-- 118
OY 297 GAGCCCACTGTGAGGAGCAATAATGTCATATGTCATTATCGCGCCACCATCGTTACT 356
Db 119 ----CCCTGTTCAGGCTGTGCTGCTATGTCATGATCGGTGCAACGGTCATCAGC 174
OY 357 ACATTTCCGGGCCATTATCAGACTGCGATGGCCGACGCCCTATGCTTATCTCTTCA 416
Db 175 ACTTGCTCAGGACCGGATATCTGATGCTCGGACAGACGCCCATGCTCATTTATCATCA 234
OY 417 ATTCTGACTTCTTCACGGGCTCATCATGCTATGCTCTCCTAATGTCATGTCCTGCTG 476
Db 235 GTTATGTAATTTGCTGCTGCGGTTGATTAATGTTGCTCTCCAAATGTCATATGTTCTGTC 294
OY 477 TTGGCAGCTTCGTAGATGATTTGGTATTTGGCTTGCTGCTGCTGCTGCTGCTTTGTAC 536
Db 295 TTTCGAGGCTTCTTAATGGTTGGTCCGGGCTCGCGGTACACTTGTCCCTGTTTAC 354
OY 537 ATTTCAAGAAATAGCCCCCTTCGAGATTAAGGTTTGTGAATACACTACACAATTCAGT 596
Db 355 ATTTCTGAACCGCTCTCCGAGATCAGAGGACAGTAAATACCTCCCTCAGTTTCTT 414
OY 597 GGATCAGAGGAATGTTCTGTCACTGATGCTGTTGGATGTCCTGTCCGCAATCA 656
Db 415 GGCTCTGGTGAATGTTTGTCACTGATGCTTTTCACTATGTCCTGAGTGACTCC 474
OY 657 CCCGATTGAGAAATTAATGCTGTTGCTGCGGATACCTTCATGTTCTTCTTGGTTTG 716
Db 475 CCTAGCTGAGAGCCATGCTCGGTGCTCTCGATCCCTTCTTCTTATTGTTCTC 534
OY 717 ACAATATTTTATCTTCTGTAATCTCCAAGATGCTGTTAGCAAGGTCGATGGCAGAG 776
Db 535 ACGGTGTTTATTTGCCCGAGTCTCTGCTGGCTGTTAGTAAAGAAAGATGACGAG 594
OY 777 GCAAAAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGTCAGGTGAATTGTCCTT 836
Db 595 GCTAAGCGATTCTTCAACAGTTATGTGCGAGAGAAAGATGTTACCGATGAGATGGCTTTA 654
OY 837 CTTCTGAAGGTTGAGGTTGAGGAGACACTTCCATTGAAGATACATCATTTGACCT 896
Db 655 CTAGTTGAAGACTAGATATAGAGAGAAAAACAATGSAAGATCTTATGTAACCTTG 714
OY 897 GCCACCGAGGACCCGATGAT--CTTGTACTGAAGGTGATAGGAACAATCACAATT 953
Db 715 GAGGATCATGAAGGTGATGATACACTTGAACCGTTGATGAGGATGACAAATGCGGCTT 774
OY 954 TATGGGCTGAAGAGGCCAGTCATGATGCTGACCTTCTAAGGGAACCCATCATGCTT 1013
Db 775 TATGAACCCACGAGAAATCAATCGTACCTTGCTAGACCTGTCCAGAACAAA----- 826
OY 1014 GGAAGTGTCTTCTCTGCACTCTGTCATGGAGCATGTGTAACGAGTGTACCCCTT 1073
Db 827 -ATAGCTCACTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATCCTT 885
```

```
OY 1074 ATGATCCGATTGTGACACTTTTGTAGTTCATGAGATATAGCTCAAGCTGAGGA 1133
Db 886 AAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGACGCGGA 945
OY 1134 AGTATGAGGACCAATGTTTCCAAACTTTGGAAGTATGTTAG-----TGTACAGAT 1187
Db 946 AACACTCGAGTGGGATTTTCCCTCATTTCCGAAGCATGTTCACTACTGCGGATGCG 1005
OY 1188 CAGCATGCCAAAAATAGCAGTGGGATGAAGAA-----ATCTTCATAGGGATGAC 1238
Db 1006 CTTACCGTAAACCGCTCATTTGGAAAAAGACATAGAGAGCCATTACAAACAACAAT 1065
OY 1239 GAGGATACGCACTGATGTTGTCAGAGGTGACTATGAG-----ACAATCTCCAT 1289
Db 1066 GATGACTATGCGACTGATGATGTTGCGGTGATGATGACTCGGACAAAGATTTGCGT 1125
OY 1290 AGCCCATTTGCTCCAGGACAGCAACAGTCCGGAAGGAAGACATTTGACCATGCT 1349
Db 1126 AGCCCTTAATGTCCGCCACAGACCAACAGCATGGA---CAAGATATGATCCCATCTT 1182
OY 1350 CACCGTGAAGTGTGTTGACATGAGAGGCAAAACCTCTTAGGGGAGGTGAGATGCT 1409
Db 1183 ACAAGTGAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCA---AGGCAACGGC 1239
OY 1410 GTGACGACACTGATATCGGTGGGATGGCAGCTTCTTGAAATGTCAGAGAAAGAA 1469
Db 1240 GAAAGTAGCATGGGAATGTTGTTGGTGGCATATGGAT----- 1279
OY 1470 GGTGAGATGTGAAAGAAAGGTGTTTCAAAAGATCTACTTGCAACCAAGAGGAT 1529
Db 1280 -----ATAGATACGAAAAAGATGAATACAGAGATATTATCTTAAGAAAGATGA 1329
OY 1530 CTTGCTCAAGAAAGGCTCAATGTTTCACTTCCCGGTGTTGCGATGTTTGAAGGT 1589
Db 1330 GCTGAATCTGCGCGTGCATCATCTATTCGGAAGTCCGATGCT-----GGA 1383
OY 1590 AGTGAATTTGTACATGCTGCTGCTTTAGTAAGTCAAGCACTTTCTCAAGGCTTT 1649
Db 1384 GGCAGCTACATTCACGCTTCTGCCCCGTGTAAGCAGATCTGTTTGTGCTCTAATCAGTT 1443
OY 1650 GCTGAACCAAGCATGTGAGATGCTGCCATGTTCAACCATCTGAGGTAGCTGCCAAAGT 1709
Db 1444 CATGA-----TCCGCATGTTTCCCGGAGAAAAATGCTGCTCTGGA 1488
OY 1710 TCAAGTTGAAAGATTTGTTGAACCTGAGTGAAGCGTGCCTGTTAGTCGCTGTTGA 1769
Db 1489 CCACTCTGCTGCTCTTCTTGAACCTGTTGAACGCTGCTGTTGTTGTTGCTCGGC 1548
OY 1770 ATTCAAGATCCTTCAACAGTTTGTGGAATAAACGCTGTTCTGFACTATACCCCAAAAT 1829
Db 1549 ATTCAATACTGCAGCAGTTTCAAGTATCAATGAGTCTCTACTACATCCTCAGATT 1608
OY 1830 CTGAGCAAGCTGCTGTCAGTATTTCTTCCAAATTGCTCAGCTCGGCATCAGCA 1889
Db 1609 CTCGAACGGGCTGCGGTAGATATTTCTTTCGAGCCTCGGACTAAGTTCATCTGCG 1668
OY 1890 TCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCTTGCAATTTGCTTGGCATG 1949
Db 1669 TCATTCCTCATCAGCGGTTTAAACAATTAATGCTCTCCAGCCATTTGCTGTCATG 1728
OY 1950 CTGCTTATGATCTTCCGGAAGAAAGTTTGTGCTGCTAGGCACAATTTCCAACTTGATA 2009
Db 1729 AGACTCATGATGATCCGGAAGAAAGTCATTAATTAATTTGGAACAATCCAGTCTCATTT 1788
OY 2010 GCATCTTAGTATCTGCTGTTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCT 2069
Db 1789 GTCTCACTTGTGCTCTTGTTCATCAGCGAGCTCATCACAATGCAAAAGTGTGAACGA 1848
OY 2070 TTGCTCTCACCGTCAAGTATATGCTACTTCTGCTGCTTCTGTTATGGGATTTGCTCC 2129
Db 1849 GCACTCTCACAGGTTGTGTGCTCTACTTCTGCTTCTTCTGATGGGTTACGGTCCC 1908
```


QY	2130	ATCCCCAACATTTTATGTGACAGAGATCTTTCCAAACCAGGGGTTGCGGCTCTGTATTTGCC	2189
Db	1909	ATTCCAAACATCCTCTGTCTGTGAATCTTCCCAACAAGATCCGGTCTCTGTCAATGCC	1968
QY	2190	ATTGTGCCTTTACATTTGGAATCGGAGATATCATCGTCACTACAGCCCTTCTGTGATG	2249
Db	1969	ATATGTGCTATGTCCTTTTGGATTTGAGACATTATGTGACGTACTCACTTCCCGTTCTC	2028
QY	2250	CTGAATGCTATTGGACTGGCGGCTGTTTTCAGCATATATGAGTTCGTATGCTGATTTCC	2309
Db	2029	CTCAGCTCGATCGGACTAGTGGTGTTTTCAGCATTTACGCTGCGGTTTGCGTTATCTCA	2088
QY	2310	TTTGTGTTGCTCTTCCCTTAAGTCCCTTGAGACAAAGGGATGCCCTTGAGGTTATTACC	2369
Db	2089	TGGATCTTCGTTTACATGAAGTCCCGGAGACTTAAGGCAATGCTTTGGAAAGTTATCACA	2148
QY	2370	GAAATCTTTGCAAGTTGGTCC	2389
Db	2149	GACTACTTTGCCCTTTGGAGC	2168

RESULT 11			
AX507559			
LOCUS	AX507559	2205 bp	DNA
DEFINITION	Sequence 2254 from Patent WO0216655.	linear	PAT 27-SEP-2002

VERSION	AX507559.1	GI:23388796
KEYWORDS	.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Harper, J.F., Kreps, J., Wang, X. and Zhu, T.	Stress-regulated genes of plants, transgenic plants containing same, and methods of use	Patent: WO 021655-A 2254 28-FEB-2002; The Scripps Research Institute (US) ; Syngenta Participations AG (CH)

FEATURES	Location/Qualifiers
source	1..2205
	/organism="Arabidopsis thaliana"
	/mol_type="Genomic DNA"
	/db_xref="taxon:3702"
BASE COUNT	521 a 524 c 521 g 639 t
ORIGIN	

Query Match	25.2%	Score 700;	DB 6;	length 2205;
Best Local Similarity	61.2%;	Pred. No. 6e-173;		
Matches 1371;	Conservative	0;	Mismatches 770;	Indels 99;
				Gaps 11;

OY		177	ATGTCGGGGCTGTTCCTTGCACCATAAGTCGCCCTCCATCGCAATTATTGACAGGGTG	236
Db		1	ATGAAGGAGCGACTCTCGTTGCTCTCGCCGCCACAATCGGCAATTTCTTACAAGATGG	60
OY		237	GACATGCCACCATCGCAGCTGCTGTCTGTATATTAAGAAGAAATTTCAATTGCAAAT	296
Db		61	GACAATGCCACCATTGCTGAGCTAATGTTATATCAACAAGAATTGAATCTACCAA--	118
OY		297	GAGCCCACGTGAGGGACTTAATTGTGTCATATGTCATTATCGGCCACCACCATCGTTACT	356
Db		119	----CCTCTGTTCAAGGCTTGTTCGTTGCTATGTCATTGAATCGGTCCAACGGTCAICACG	174
OY		357	ACATTTCTCCGGGCCATTATTCAGACTCGATTGGCCGACGCCCTATGCTTATTCCTCTTCA	416
Db		175	ACTTGCTCAGGACCGATATTCGATTGGCTCGGACGAGCCCATGCTCATTTTATCATCA	234
OY		417	ATTCTGTACTTCTTCAAGCGGCTCATCATGCTATGTTCTCCCTAATGTCATGTCCTGCTG	476
Db		235	GTTATGTAATTCGTCCTCGGTTTGATTAATGTTGTGTTCTCCCAATGTCATATGTTCTGTGC	294
OY		477	TTGCGACGCTTCGTAGATGAATTTGGTATTTGGCTTGGCTGTCAACGCTGTGCTTTGTAC	536

Db 295 TTGCTAGGCTTCTTAATGGGTTTGGTGCCGGGCTCGCGGTACACTTGTCCCTGTTTAC 354

QY 537 ATTCAGAATAAGCCCCCTTCGAGATTAGAGTTTGCTGAATACACTACCAATTCAGT 596

Db 355 ATTTCTGAACCCGCTCTCCGGAGATCAGAGSACAGTTAATACTCTCCCTCAGTTTCTT 414

QY 597 GGATCAGAGGAATGTTCTGTCTACTCTGCATGCTGTCTTGGGATGTCCCTGTCCCATCA 656

Db 415 GGCTCTGGTGGAATGTTTTTGTCTACTGTATGTTTTCACTATGTCCCTGAGTGACTCC 474

QY 657 CCCGATTGGAGAATTATGCTTGGTGTGCTCCGCAATCTCATTTGTTCTTGGTTTTG 716

Db 475 CCTAGCTGGAGAGCCCATGCTCGGTGTCTCTCCATTCCTTCTCTTCTTAATTTGTTTTCTC 534

QY 717 ACAATATTTTATCTTCTCGAATCTCCAGATGGCTGTGTTAGCAAAGGTCGGATGSCAGAG 776

Db 535 ACGGTGTTTATTTTGCCCGAGTCTCTCGTTGGCTGCTTAAGTAAAGAGAAGATGACGAG 594

QY 777 GCAAAAAGTGTTGCAAAAGTTACGGGGGAAGAATCTCAGGTGAATTGCCCTT 836
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GCTAAGCGAGTCTTCAACAGTTATGTGCAGAGAGATGTTACCGATGAGATGGCTTTA 654

837 CTTCGAGCGTTGACGTTGAGGAGACACTTCCATTGAAGATACATCATTTGCACT 896

897 GCCACCGAGCAGCCCATGAT---CTTGTACTGACGGTGATAAGGACAACAATCACTT 953

Db 715 GAGGATCATGAAGTGATGATACACTTGAACCGTTGATGAGGATGGACCAATGCGGCTT 774

y 534 TATGGGCTTGAAGAAAGCCAGTCATGGATGCCTGACCTCTTAAGGAGCCCACTCATCCTTT
 |||||
 Db 775 TATGGAACCCACGAGATCATCTGTACTCTGTAGACCTGTCCAGAACA----- 826

QY	1014	GGAAGTGTGCTTCTCTTGATCTCGTCATGGAGCATGGTGAACCGAGTGTACCCCTT	1073
Db	827	-ATAGCTACCTTGGCGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATCTT	885

QY 1074 ATGATCCGATTGTGACATTCTTTGGTAGTGTCCATGAGATATGCCCTCAAGCTGAGGA 1133

Db 886 AAAGATCCGCTGTCATCTTTTGGCAGTCTCCACGAGAAGATGCCAAGACGCGGA 945

QY 1134 AGTATGAGGACACATTTGTTCCAAACTTTGGAAGTATGTTCA-----TGTCAAGAT 1187

Db 946 AACACTCGAGTGGGATTTTCCCTCATTTCCGAGCATGTTCACTACTGCCGATGCG 1005

Db 1006 CCTCAGGTAACCGGCTCATTTGGAAAAGACATAGAGAGCCATTACACAAAGACAAT 1065

```
QY      1239 GAGGAGTACGCATCTGATGGTCGAGAGGTGACTATGAGS-----ACAATCTCCAT   1289
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1066 GATGACTATGCCACTGATGATGGTGCGGTGATGATGATGACTCGACAACGATTTCGT   1125
```

QY 1290 AGCCATTGCTGTCCAGGCAGGCACAGGTGCGAAGGAAGGACATTGTGCACCATGGT 1349

DB 1126 AGCCCCCTAAATGTCGGCAGACCACCAAGCAATGGA---CAAGGATATGATCCACATCC 1182

DY 1350 CACCCTGGAAGTCCTTTGAGCATGAGAAGGCAAACCTCTTAAAGGAGGGGTGAGATGCT 1409

Db 1183 ACAAGTGAAGCACTTTAAGCATGAGACGACACAGTAACCTTATGCA--AGGCAACGGC 1239

Db 1240 GAAAGTAGCATGGGAATTGGTGTGTGGCATATGGGAT----- 1279

```

QY      1470 GGTGAGATGCTAGAAAGCAAGGTGTTCCAAGAAGCTCTACTTGCACCAAGAGGGAGTT 1529
          |||||
Db      1280 -----ATAGATACGAAACGATGAATACAGAGGTATTATCTTAAAGAGATGGA 1329

```

QY 1530 CCTGGCTCAGAGGGGCTCAATTGTTTCACTTCCCGGTGGTGGCGATGTTTGAAGGT 1589
||| ||| | ||||| | | ||||| ||| ||||| |

Db 1330 GCTGAATCTCCCGCTGGCTCGATCATCTTAATCCCGAGGTCGGATGGT-----GGA 1383

Qy 1590 AGTGAATTGTGATGCTGCTGCTTTAGTAAGTCACTCAGCAGCTTTCTCAAGGGCTCTT 1649

Db 1384 GGCAGCTACATTCACGCTTCTGCCCTGTGAAGACAGATCTGTTCTTGCTTAATCAGTT 1443

Qy 1650 GCTGAACCAAGCATGTAGATGCTGCCATGCTTCAACCCATCTGAGGTAGCTGCCAAGGT 1709

Db 1444 CATGA-----TCCGCCATGGTTCCCGGAGAAAATGCTGCTCTGGA 1488

Qy 1710 TCAGTTGGAAGAATTTGTTGAACCTGAGTGAGCGCTGCCCTGTAGTCCGCTTGA 1769

Db 1489 CCACTCTGGTCTGCTCTTCTGAACCTGGGTGAAGCGCTGCCCTGGTGTGCTGCC 1548

Qy 1770 ATTCAAGTCTTCAACAGTTTGTCTGAATAACGGTGTCTGTACTATACCCAAAATT 1829

Db 1549 ATTCAATACAGCAGCAGATTTCAGGTATCAATGAGTTCTCTACTACACTCTCAGATT 1608

Qy 1830 CTGAGCAGCTGGTGTGAGATTATCTTCCAAATTGGTCTCAGCTCGGCATCAGCA 1889

Db 1609 CTGGAACGGGCTGGCGTAGATATTCTTTGAGCCTCGGACTAAGTTCCATCTGCG 1668

Qy 1890 TCCATCTTGATGAGTTCTCTCACTACCTTAATGCTTCTGCTTGAATGGCTTGGCAATG 1949

Db 1669 TCATCTCTCAACAGCGGTTTAAACAATTAATCATGCTCCAGCCATGTGCTGGCATG 1728

Qy 1950 CTGCTTATGATCTTTCCGGAAGAGTTTGTGCTGCTAGCAGCAATTCATCTTGATA 2009

Db 1729 AGACTCAATGATGATCCGGAAGAGGTATTAATCTTCTGAGCAATCCAGTCTCAT 1788

Qy 2010 GCATCTTAGTTATCCTGCTGTGTCCAATCTAATGATTGGGTACATAAGCCCATGCT 2069

Db 1789 GTCTCACTGTGCTGCTCTGTATCAGCAGCTCATCCACATCAGCAAAAGTCTGAACGCA 1848

Qy 2070 TTGCTCTCAACCGTCAAGTGTATCGTCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2129

Db 1849 GCATCTTCCACAGGTTGTGTGCTGCTTACTTCTGCTTCTGCTGCTGCTGCTGCTGCT 1908

Qy 2130 ATCCCAATTTTATGTGAGAGATCTTCCAAACAGGCTTGGCTCTGCTGCTGCTGCTGCT 2189

Db 1909 ATTCCAACATCTCTGTTCTGAATCTTCCCAACAGAGTCCGTGCTCTGCTGCTGCTGCT 1968

Qy 2190 ATTGCTCTTATCATTTCTGATCGAGATCATCTGCTACCTACAGCTTCTGCTGCTGCTGCT 2249

Db 1969 ATATGCTGATGCTCTTTGATGAGACATTAATGTCACGTACTCACTTCCGCTCTC 2028

Qy 2250 CTGAATGCTATGAGTGGCGGTGTTTACGATATATGACATATGACAGTCTGATTTCC 2309

Db 2029 CTCAGCTGATCGGACTAGTGTGTGTTTCAACATTTACGCTGCGGTTTCCGTTATCTCA 2088

Qy 2310 TTTGCTGTGCTCTTCTTAAGGTCCTTGAAGCAAAAGGGGATGCCCTTGAAGTTATTACC 2369

Db 2089 TGGATCTTCTGTTTACATGAAAGTCCCGAGACTAAAGGCATGCTTTGGAAGTTATCACA 2148

Qy 2370 GAATCTTTGCAAGTGTGC 2389

Db 2149 GACTACTTGTCTTTGAGC 2168

RESULT 12

LOCUS AR208571 1487 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 13 from patent US 6383776.

ACCESSION AR208571

VERSION AR208571.1 GI:21509759

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1487)

AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.

TITLE Plant sugar transport proteins

JOURNAL Patent: US 6383776-A 13 07-MAY-2002;

FEATURES

Source 1.1487

BASE COUNT 355 a 320 c 373 g 439 t

ORIGIN

Query Match 24.9%; Score 691.2; DB 6; Length 1487;

Best Local Similarity 80.1%; Pred.No.1.2e-170;

Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

Qy 1395 GAGGTGAGAGATGCTGTGAGCAGCACTGATATCGGTGGGATGGCAGCTTGCTGGAAA 1454

Db 12 GAGGGTGGGAGGAGCAGTACAGCAGCACTGATTTGGTGGGGGTGGCAACTCGCATGAAA 71

Qy 1455 TGGTCAAGAGAAGAGGTGAGATGAGTAAGAAAGAGAGGTGTTTCAAAAGAGTCTACTTG 1514

Db 72 TGGTGGAGCGACAAAGCGAGGATGGCAAGAAAGAGAGGCTTCAAAAGATCTACTTG 131

Qy 1515 CACCAAGAGGAGTTCCTGGCTCAAGAAAGGGCTCAATGTTTCACTTCCCGTGGTGC 1574

Db 132 CACCAAGAGGAGGTGGCGGACTCAAGAAAGGGCTCTGTTGTTTCACTTCTGTTGGGGT 191

Qy 1575 GATG--TTTGGAGGATGAGTTTGTACATGCTGCTGCTTAAAGTACAGTACGA 1631

Db 192 GATGCCACGCAAGGGGCGAGTGGGTATATACATGCTGCTGCTTGGTAAGCCACTCGGCT 251

Qy 1632 CTTTCTCAAGAGGTCTGCTGAACCAAGCATGTACAGATCT--GCCATGTTACCCA 1688

Db 252 CTTTACTCCAAGGATCTTATGGAAGAGCGTATGGCGCGGCTCCAGCCATGATTCATCCA 311

Qy 1689 TCTGAGTAGCTGCCAAGGTTCAAGTTGGAAGATTGTTGAACCTGAGTGAAGCGCT 1748

Db 312 TTGAGGAGAGCTCCCAAGGTTCAATCTGAAAAGATCTGTTGAACCTGCTGAGAGCGT 371

Qy 1749 GCCCTGTTAGTGGGTGTTGAATTCAGATCTCTCAACAGTTGCTGGAATAAACGGTGT 1808

Db 372 GCATTTGCTGCGGTGTTGGAATTCAGATCTTCAAGAGTTGCTGGAATAAATGAGATT 431

Qy 1809 CTGTAATATACCCCAAAATCTTGAAGCAAGCTGAGTGGCAGTTATCTTCCAAATTT 1868

Db 432 CTCTACTATACTCTCAATTCGAGAGCAAGCTGAGTGGCTGTTCTTCCATCTT 491

Qy 1869 GGTCTACGCTGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACTACTATGCTT 1928

Db 492 GGCTCAGTTACGATCAGCATCCATCTTGATCAGTTCTCTCACTACTACTGCTC 551

Qy 1929 CCTTGCATTTGGCTTTGCCATGCTCTTATGATCTTTCCGGAAGAGGTTTGGCTGCTA 1988

Db 552 CCAAGCATTTGGTGTAGCCATGAGACTTATGATATATCTGGAAGAGGTTTCTGCTACTG 611

Qy 1989 GGCACAATTCGAATCTTGAAGATCTCTAGTATCTGTTGTGCCAATCTAATTGAT 2048

Db 612 GGCAAAATCCCATCTTGAATGATCCCTAATGTTTGGGTGTGTCATATGATCAAC 671

Qy 2049 TTGGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAAGTTATCGTACTTCTGCTGC 2108

Db 672 TTGAGTACGGTGCCCAAGCTGTGCTCTCCACAGTTAGCTCATGTCTACTTCTGCTGC 731

Qy 2109 TTGTTATGGGATTTGTGCCATCCCAACATTTTATGTGAGAGATCTTCCAAACGAG 2168

Db 732 TTTGTATGGGCTTTGGCCCATCCCAACATTTATGTGAGAGATTTTCCCAACGAGA 791

Qy 2169 GTTCGTGGCTCTGTATTTGCTGCTTACATTTCTGATCGGAGATATCATGCTC 2228

Db 792 GTCCGTGGTGTCTGATCGCTATTTTGGCCCTCACATTTTGTGACATTAATGTT 851

Qy 2229 ACCTACAGCCTTCTGTGATGCTGAATGCTATTGAGTGGCGGGTGTTCAGCATATAT 2288

Db 852 ACCTACAGCCTGCTGTGATGCTGAATGCTATTGTTAGCGGGGTCTTTGTATATAT 911

Qy 2289 GCAGTGTATGCTTGAATTTCTTGTGTTGCTTCTTAAAGTCCCTGAGACAAAGGG 2348

Db 912 GCAGTGTGTTGCTGATGCTTGTGTGTTGTTCTACTTAAAGGTCCAGAGACAAAGGG 971

QY	2349	ATGCCCCCTTGAGGTATTACCGAATCTTTGACAGTTGGTGGCAAGCAGC--GGCTGCA	2405
Db	972	ATGCCCCCTGAGGTCATCACCGAGTCTTTGCGGTTGGGCGAAGCAGCGACCAACC	1031
QY	2406	AAAGCCTAATTTCTTTGGTACCTTTGTGCAACTATTGCACTG	2449
Db	1032	ATGCGCTGATTTCATCATGAGAGCTTTGTTTTCAGTTTGCAACTG	1075
RESULT 13			
ATSUGTRPR			
LOCUS	ATSUGTRPR	2426 bp	linear
DEFINITION	A.thaliana mRNA for sugar transporter.		
ACCESSION	250752		
VERSION	250752.1	GI:1495272	
KEYWORDS	sugar transporter.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 2426)		
TITLE	Tjaden,J. and Neuhaus,E.		
JOURNAL	A new sugar transporter protein from Arabidopsis thaliana		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 2426)		
TITLE	Neuhaus,E.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (09-AUG-1995) Neuhaus E., University of Osnabrueck, Pflanzenphysiologie, Barbarastr. 11, Osnabrueck, Germany, 49069		
FEATURES	location/Qualifiers		
SOURCE	1..2426		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	37..2241		
	/codon_start=1		
	/product="sugar transporter"		
	/protein_id="CAA90628.1"		
	/db_xref="GI:1495273"		
	/translation="SPTREMBL:Q96290"		
	/translation="MKGATLVALATIGNFLQGMDNAITAGAMVYINKDNLPTSVOG		
	LVVAMSLIGATVITTCSPISDMLGRPMILISSVMYFVCGLIMLWSPNYVLCFARL		
	LNGFGAGLAVTLVPVYISETAPEIRGQNLTPQLFGSGGMFLSYCMVFTMSLSDSPS		
	WRAMLGVLISPLVFLVTVFYLPESPRMLVSKRMDKRVLOQLCGEDVTDEMAL		
	LVEGLDIGEKTMEDLLVLTLEDHGGDTLETVEDGQIRLYGTHENOSYLARPYEQN		
	SLGLRSRHGSLANOSMILKDPVLNLFGLHEKYPEAGNTRSGIFPHGSMESTTAD		
	APHGKPAHWEKDIESHYNNKNDNDYATDDGAGDDDDSDNDLRPLMSRQTTSMDKMIP		
	HPTSGSTLSMRHSTLMQNGESSMIGCGMHWGYRYENDEYKRYLKEGDAESRRGS		
	IISIPGGPDGGSYIHASALVSRVLPKSVHGSAMVPEKIAASGPLMSALLEPGVK		
	RALVVGVGIIQLQOFSGINGVLYTTPQLERAGVDILSLGLSSISAPLISGLTTL		
	LMLPAIVAMRLMDVSGRSLILMTIPLVLSVVLVISELIHISKVNAALSTGCVV		
	LYCFEPMVGYPQOTSSVLKSSQADRGLCAICAMVFWIGDIIVTYSLPVLLSIEL		
	VGVFSIYAACVLSWIFVYMKVPETKMPLEVTIDYFAFGAQAQASAPSKDI"		
BASE COUNT	577 a	569 c	544 g 735 t 1 others
ORIGIN			
Query Match	24.2%;	Score 672.2;	DB 8; length 2426;
Best local Similarity	61.0%;	Pred. No. 1.3e-165;	
Matches 1367;	Conservative 0;	Mismatches 773;	Indels 101; Gaps 13;
QY	177	ATGTGGGGGGCTGTTCTTGTCGCCATAGTCGCCCTCCATCGCAATCTATTGACAGGGGTGG	236
Db	37	ATGAAGGAGCGACTCTCGTTGCTCTCGCCGCCACAATCGGCAATTTCTTACAGGATGG	96
QY	237	GACAATGCCACCATCGACGCTGCTGTCTGTATATAAAGAAATTTCAATTGCAAAAT	296
Db	97	GACATATGCCACCATGTGCTGAGAGTATGTTATATCAACAAGAAGCTGAATCTACAA--	154
QY	297	GACCCCACTGTGAGGAGCTAATTTGTCAATGTCACCTATCCGCCACCATCGTTACT	356
Db	155	---CCTCTGTTCAAGGCTTGTGCTGTTGCTATGTCATGATCGGTGCAACGGTATCAACG	210

QY	357	ACATTCCTCCGGGCCATTATCAGACTCGATTGCGCGACGCCCTATGCTTATCTCTCTTCA	416
Db	211	ACTTGCTCAGGACCAGATATCTGATTGGCTCGGCAGACGCCCATGCTCATTTTATCATCA	270
QY	417	ATTCTGTACTTCTTCAGCGGCTCATCATGCTATGCTCTCTAATGCTATGCTCTG	476
Db	271	GTTATGTAATTCGTCTGCGGGTTTGATTAATGTTGCTCTCCCAATGCTATGTTCTGTGC	330
QY	477	TTGGCAGCCTTCGTAGATGGAATTTGGTATTTGGCTTGGCTGTCAAGCTTGCTTTGTAC	536
Db	331	TTTGCTAGGCTTCTTAATGGGTTTGGTGCCGGGCTCGCGGTTACACTTGTCCTGTTTAC	390
QY	537	ATTTCAGAAATAGCCCCCTTCGGAGATTAGAGTTTGCTGAATACACTACCAATTCACT	596
Db	391	ATTCTGAACCGCTCTCTCCGAGATCAGAGGACAGTTAATACTCTCCTCAGTTTCTT	450
QY	597	GGATCAGAGGAATGTTCTGTCTACTACTGCATGCTGTTGGATGTCCTGTCGCATCA	656
Db	451	GGCTCTGTGGAATGTTTGTGCATACTGTATGCTTTCACTATGTCCTGAGTGACTCC	510
QY	657	CCCGATTGAGAAATATGCTTGTGCTGCTCGCATACCTTCATTTGTTCTTGTGTTG	716
Db	511	CCTAGCTGAGAGCCATGCTCGGTCTCTCGATCCCTCTCTTATTTGTTCTC	570
QY	717	ACAATATTTTATCTTCTCTGAATCTCCAGATGCGTCTTACCAAGGTCGATGGCAGAG	776
Db	571	ACGGTGTTTTATTTGCTCCGAGTCTCCTCGTTGGCTGTTAGTAAAGGAATGGAACGAG	630
QY	777	GCAAAAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGCTCAGGTGAATTGCTCCTT	836
Db	631	GCTAAGCAGTCTTCAACAGTTATGTGCGAGAAAGATGTTACCGATGATGGCTTTA	690
QY	837	CTTCTCGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGATACATCATTTGAACT	896
Db	691	CTAGTTGAAGACCTAGATATAGAGAGAAACAAATGGAAGATCTCTTAGTAATTG	750
QY	897	GCCACCAGGCAAGCCGATGAT--CTTGTTACTGACGGTGATAAGAAACAATCACACTT	953
Db	751	GAGGATCATGAAGGTGATGATTAACACTTGAACCGTTGATGAGATGGAACAATAACGGCTT	810
QY	954	TATGGGCTGAAGAGGCCAGTCATGATTTGCTTGAACCTTCTAAGGGAACCATCATGCTT	1013
Db	811	TATGAACCCACGAGATCAATCGTACCTTGCTAGACCTGTCCAGAACAA-----	862
QY	1014	GGAAGTGTCTTCTCTTGCACTCTGTCATGAGGACATGTTGAACACAGTGTACCCCTT	1073
Db	863	-ATAGCTCACTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATCCTT	921
QY	1074	ATGATCCGATTTGTACACTTTTGTAGTGTCCATGAGATATGCTCAAGCTGAGGA	1133
Db	922	AAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAAGACAGGCGGA	981
QY	1134	AGTATGAGGACCACTTGTTCCTCAAACTTTGGAAGTATGTTTCAG-----TGTCAAGAT	1187
Db	982	AACACTCGAGTGGGATTTTCCCTCATTTCCGAAGCATGTTCACTACTGCGGATGCG	1041
QY	1188	CAGCATGCCAAAAATGAGCAGTGGGATGAAGAG-----ATCTTCAATAGGGATGAC	1238
Db	1042	CCTCACGTTAAACGGGCTCATTTGGAAAAGACATAGAGGCCATTACAAACAAGACAAT	1101
QY	1239	GAGAGTACGCACTGATGTTGTCAGAGGTTGACTATGAGG-----ACAATCTCCAT	1289
Db	1102	GATGACTATGCGACTGATGATGTTGCGGGTATGATGACTCGGACAACGATTTGCGT	1161
QY	1290	AGCCCATTTGCTTTCAGGACGGACACAGGTCCGAAGGAAGACATTTGTGACCATGCT	1349
Db	1162	AGCCCCTTAATGTGCGGCCAGACCAACAGCATGGA---CAAGATATGATCCACATCCT	1218
QY	1350	CACCGTGAAGTCTTTGAGCATGAGAAAGCAAACTCTTTAGGGGAGGGTGGAGATGCT	1409
Db	1219	ACAAGTGAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCA---AGGCAACGGC	1275


```

REFERENCE 1 (bases 1 to 135406)
AUTHORS Bhargava,A., Dalal,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M.,
Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K.,
Mohapatra,T. and Singh,N.K.
TITLE Complete sequence for Oryza sativa chromosome 11 Clone
OSJNBa0094M06
JOURNAL Unpublished
AUTHORS
REFERENCE 2 (bases 1 to 135406)
AUTHORS Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshabu,K.,
Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bhargava,A.,
Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2002) IIRGS, NRC on Plant Biotechnology, Indian
Agricultural Research Institute, LBS Centre, New Delhi, Delhi
110012, India
REFERENCE 3 (bases 1 to 135406)
AUTHORS Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshabu,K.,
Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bhargava,A.,
Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2003) IIRGS, NRC on Plant Biotechnology, Indian
Agricultural Research Institute, LBS Centre, New Delhi, Delhi
110012, India
REFERENCE 4 (bases 1 to 135406)
AUTHORS Bhargava,A., Dalal,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M.,
Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K.,
Mohapatra,T. and Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2003) IIRGS, NRC on Plant Biotechnology, Indian
Agricultural Research Institute, LBS Centre, New Delhi, Delhi
110012, India
COMMENT On Jun 6, 2003 this sequence version replaced gi:28604240.
FEATURES
source location/Qualifiers
1..135406
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="11"
/clone="OSJNBa0094M06"
BASE COUNT 39875 a 28472 c 27393 g 39665 t 1 others
ORIGIN
Query Match 23.0%; Score 637.8; DB 8; Length 135406;
Best Local Similarity 56.6%; Pred. No. 2.1e-156;
Matches 1327; Conservative 0; Mismatches 972; Indels 46; Gaps 6;
QY 84 CCGTCAGTTGTTCTTCTTGGCTTGGCCCTGCGCTTCCGCTGCTTCTTCTTACAGAG 143
DB 99067 CTGTTGGTTGCAAAATCATCCAAATCCCTTCGCGCTTTTTCATGTTGTTCTGAGA 99126
QY 144 CCGGTGACCTCGACGATATCTTGAGAGACAGATGTCGGGGCTGTTCTTGTGCCATA 203
DB 99127 AACCGATCCGAGTCGATCGATCGCACGTAAGATGAGAGCGCGCTGTTGTGGCGTG 99186
QY 204 GTGCGCTCCATCGGCAATCTATTGCAAGGGGTGGACAATGCCACCATGCACTGCTTT 263
DB 99187 GCGCGCGCATGCGCACTACCTGCAAGGATGGACAACGCCACCATGCGCGCGCTG 99246
QY 264 CTGTATATAAAGAAATTTCAATTCGAAATGAGCCCACTGTGAGGGAATAATTGTG 323
DB 99247 CTGTACATCAAGCGGAGTTGCGCCCTCGAGACCCAGCCCGCTGAGAGGCTCTGTC 99306
QY 324 TCAATGTCATATTCGCGCGCCACCATCGTTACTATCTCCGGGCCATATCAGACTCG 383
DB 99307 GCCATGTCCTCATTCGCGCGCCACCATCATCACTTCTCCGGCCCTCTCCGACCTC 99366
QY 384 ATTGCGCGAGCGCCCTATGCTTATTTCTTCAATTTCTGTAATTTCTTCAAGCGGCTCATC 443
DB 99367 GTGCGCGCGCGCCCATGCTCATGCGCTCTCTCTCTCTTACTTCCGCGCGCTCATC 99426
QY 444 ATGCTATGTTCTCTAATGTCTATGTCTGCTGTTGGACGCTTCGTAGATGATTTGGT 503

```

```

DB 99427 ATGCTCTGTGTCGCCCAAGCTCTACGTCTCTCTCTCGCCCGCTCGTCAAGGCTCGGC 99486
QY 504 ATTGCTTGCTGTCAAGCTTGTGCTTGTGATCATTTGAGAATAAGCCCTTGGAGATT 563
DB 99487 GTGCGCTCGCGCTACACCTCTGTCCTCCGCTTACATCTCCGAGACCTCCCGCGAGATC 99546
QY 564 AGAGTTTGTGTAATACCTACCAATTCAGTGTGATGAGGAAATGTTCTTGTATAC 623
DB 99547 CGCGCGCGCTCAACACGCTGCGCAGTTACCGCGCTCGCGCGCATGTTATGTTCTTAC 99606
QY 624 TGCATGTTGTTGGAGTCTCTGCTGCGCCATCACCCGATGGAGAATTATGCTTGTGTG 683
DB 99607 TGCATGATCTTCCCATGACGCTCTCGCCCTCGCCCACTGCGCATCATGCTGCGCTC 99666
QY 684 CTCGCGATACCTTCATGTTCTTCTTGTGTTGACAAATTTATCTTCTGAATCTCCA 743
DB 99667 CTCTGCTCCCTCGCTGCTCTACCTGTTGTCACCGCTCTTCTACCTCCCGAGTCCCG 99726
QY 744 AGATGCTCTGTTAGCAAAAGCTGGATGCGACAGGCGCAAAAAGCTTGCAGAAAGTTACG 803
DB 99727 CCGTGGCTCTGTCAGCAAGGGCGCATGAAGAGGCGCAGGTTGTGCTCGAGATGCTGCG 99786
QY 804 GGGAAAGACGATGTTCTCAGTGAATTTGCTTCTTCTTCTGGAAGGTTGAGGTTGAGGA 863
DB 99787 GCGCGGAGGACGTTCTCGGAGATGCGCTGCTGTTGAGGGCTCTGCGACCGCGCGC 99846
QY 864 GACACTTCCATTTGAAGATCATCATTTGACCTGCGCACCGAGCGACCGATGATTTGTT 923
DB 99847 GACACGAGATCGAGACTGCTGCTGCGCCCTCGGCGCTCGAGGCGACCGCGCA-----G 99900
QY 924 ACTGACGTGATTAAGAAACAATCACACTTATGGGCTGGAAGAGGCCAGTCATGAT 983
DB 99901 AACGAGCAGCGGAGGACACCGTCAACGCTGTACGGCGCGGAGCAGGGCTTCTGTTGGT 99960
QY 984 GCTGACCTTCTTAAGGAGCCCATCATGCTTGAAGTGTCTTCTTGCATCTGTCAT 1043
DB 99961 GCGCAGCGGCTCGCGCGCGCGCGGACGATGCTGGGAGCTCGCTGGGCTGCAAGCG 100020
QY 1044 GGGAGCATGTTGAACCAAGATGTACCCCTTATGATTCGATTTGACACTTTTGTGAT 1103
DB 100021 TCGCGCATGCGACGATGTACGACGATGAAGGACCCGCTGTGTGCGCTTCTCGGAGC 100080
QY 1104 GTCCATGAGATATGCTCTCAAGCTGAG-----GAAGTATGAGGAGCA 1148
DB 100081 GTCCAGAGCGGCTGCGCGAGTCCGCGCGCGCGCGCCAGCCAGCATGAGGGGAGCAG 100140
QY 1149 TTGTTTCCAACTTGAAGATGTTCAAGTGTACAGATCAGC----- 1191
DB 100141 CTGTTCCTCAACCTCGGAGCATGCTTAAGCTCAACGATAGCCCGCGGACAGACTGG 100200
QY 1192 -ATGCCAAATAGACAGTGGATGAAGAAATCTTCATAGGATGACGAGAGTACGA 1250
DB 100201 GACGAGAGAAACGTGCAAGCTTGGCAGCAGCACTTCGAGAGGAGGAGGAGTACTTC 100260
QY 1251 TCTGATGTTGACGAGGATGATTAAGACAATCTTCATAGCCCATTTGCTGTCAGGAG 1310
DB 100261 TCCGACGACGCAAGGACGACGACGCGCGGCTGCGAGGCAACGCTGTGTGCGG 100320
QY 1311 GCAACAGTGCAGGAGGAGGACATTTGACCAATGTTCAACCTGGAAGTGTCTTGAGC 1370
DB 100321 CAGAG---GACCGAGCTGAGACCAAGAACGAGCCCGCTTCGCGCAGGTGCGATGAG 100377
QY 1371 ATGAGAGGCAAAACCTCTTAAGGAGGAGGTGAGATGTTGACGACACTGATATCGGT 1430
DB 100378 CGGACAGACGATCGCGCGCGCGCGCGCTGAGACGCGGACACCATGCGCATCGCG 100437
QY 1431 GGGGATGCGACGTTGCTTGAATAATGCTCAGAGAGAGAGTGAAGATGTAGAAAGAA 1490
DB 100438 GCGCGGTGGCACTGCGCTGGAAGTGAACGAGAGACGTGGGCCCGACGCGCTCAAGCGC 100497
QY 1491 GGTGTTTCAAAAGAGTCTACTTTCACCAAGAGGAGTTCTGCTCAAGAGGGGCTCA 1550

```


Db 100498 GGGCGGTGAAGCGCATGTACCTGCACGAGAGTCCAGGCCGCCGCCGGGTGACTCG 100557
QY 1551 ATTGTTCA--CTTCCCGGTGTGGCCGATGTTTTTGAGGGTAGTGAGTTGTACATGCTG 1608
Db 100558 GGGCGCGCGCGCGCAGCGCGCAGTCCACGGCGGTACGTCCACGGCGCGCGCTGTGAGCCGG 100617
QY 1609 CTGCTTTAGTAAGTCAAGTCAAGCACTTTTCAAAAGGGTCTTGCTGAACCAAGCATGTGAG 1668
Db 100618 TCGATGCTGTACCAAGAGAGCTCTGATCGGGCAGAGACCCCAAGAGCCGGCGTTCCGG 100677
QY 1669 ATGCTGCCATGCTTCAACCATCTGAGTAGCTGCCAAAGGTTACGTTGAAAGATTGT 1728
Db 100678 AACCGGCC--GAGGCGGTGGCGCGCGCGGTGACGGGCCCGCGGTGGCGGAGCTGC 100735
QY 1729 TTGAACCTGAGTGAAGCGCGTCCCTGTAGTGGGTGTGAATTGAGATCCTTCAACAGT 1788
Db 100736 TCGAGCGCGGGGTCCGCCACGCCCTCTTGCGGGCGGTACACCATCCAGATCCTCCAGCAAT 100795
QY 1789 TTGCTGAATAAACGGTGTCTGTACTATACCCCAAAATTCCTGAGCAAGCTGTGTGG 1848
Db 100796 TCTCCGGCATCAACGGCGGTGTCTACTATACACCCCGAGATCCTCGACCAAGGCCGGCTCA 100855
QY 1849 CAGTTATCTTTCCAAATTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTC 1908
Db 100856 GCGTCTCTCGCCAGCCTCGGCCCTCTCCGGCGACTCCACCTTCATCTCATCAAGCGGCC 100915
QY 1909 TCACCTACCTTACTATGCTTCTTGCATGAGCTTGCATGCTGTATGATCTTCCG 1968
Db 100916 TCACCAAGCTTCTCATGCTCTCCGTCATCGCGGTGCGCATGCGCTCATGAGCGCTCGG 100975
QY 1969 GAAGAAGTTTGTGCTGCTAGGACAAATTCATCTTGATAGCATCTAGTTATCCTGG 2028
Db 100976 GCCGCGCGCCTCTCTCTGTGACGCTGCGCTCTGTCGCGTCCCTGCGCGTCTCG 101035
QY 2029 TTGTGCAATCTAATTTGATTTGGGTACACTAGCCCATGCTTGTCTCCACCGTCAAGT 2088
Db 101036 TGGTGGCGAAGTGTGCGCCATGCGCGCGCAACCGCGCGCTGTGACGGGAGCG 101095
QY 2089 TTATCGTCTACTTCTGCTGCTTGTATGAGATTGTGCCATCCCAACATTTTATGTG 2148
Db 101096 TCATGCTTACTTCTGCTGCTTGTATGAGGTTTGGCCCCATCCCAACATCTCTGCG 101155
QY 2149 CAGAGATCTTCCCAACCAAGGTTGCTGCTCTGTATTTGCCATTTGCTTACATCT 2208
Db 101156 CCGAGATCTTCCGACGAGGTTGAGGGGACTCTGCAATTGCCATCTGCTGAGCGTTCT 101215
QY 2209 GGATCGAGATATCATGCTCACCTACAGCTTCTGTGATGCTGAATGCTATTTGACTGG 2268
Db 101216 GGTGCGGACATCGCGCTCACGTACAGCTTCCCGTCAATGCTCAAGCTCCGTGGGCTCG 101275
QY 2269 CGGTGTTTCAGCATATATGAGTGTGATTTCTTGTGTTGCTTCTCTTA 2328
Db 101276 CCGCGGTGTCTCTCTTACGCGCGGTGTGCGCTGCGCTGCTGCGCTCA 101335
QY 2329 AGTCCCTGAGACAAGGGGATGCCCTTGAAGTTATTACCGAATCTTTGCAGTGTG 2388
Db 101336 AGTGCCCGAGACCAAGGGCTCCCGCTCGAGGTATCATGAGTCTTCAACGTGCGCG 101395
QY 2389 CGAAG 2393
Db 101396 CCAAG 101400

Search completed: January 5, 2004, 14:01:37
Job time : 10118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 09:46:20 ; Search time 729 Seconds
(without alignments)
10283.044 Million cell updates/sec

Title: US-10-051-909-31

Perfect score: 2777
Sequence: 1 cacggggttagctcgagc.....aaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200.2	79.2	2824	24	ABK51962
2	2200.2	79.2	2824	25	ABX93198
3	940.2	33.9	2601	24	ABK51965
4	940.2	33.9	2601	25	ABX93201
5	875.8	31.5	2190	24	ABZ13510
6	700	25.2	2205	24	ABZ14449
7	691.2	24.9	1487	24	ABK51968
8	691.2	24.9	1487	25	ABX93204

9	594.6	21.4	1692	24	ABK51966
10	594.6	21.4	1692	25	ABX93202
11	394.4	14.2	1009	24	ABK51969
12	394.4	14.2	1009	25	ABX93205
13	257.2	9.3	282	24	ABL71325
14	227	8.2	870	24	ABK51964
15	227	8.2	870	25	ABX93200
16	143.4	5.2	155	25	ABX86137
17	123.2	4.4	510	24	ABK51967
18	123.2	4.4	510	25	ABX93203
19	118.6	4.3	903	22	AAC83232
20	114.2	4.1	443	24	ABK51963
21	114.2	4.1	443	25	ABX93199
22	112.4	4.0	299	24	ABL76720
23	96.2	3.5	1853	24	ABK51973
24	96.2	3.5	1853	25	ABX93209
25	87.6	3.2	1826	21	AAC42332
26	84.8	3.1	1527	21	AAC43261
27	84.8	3.1	1662	24	ABL41878
28	84	3.0	2017	24	ABK51972
29	84	3.0	2017	25	ABX93208
30	82.8	3.0	914	21	AAC44159
31	81.4	2.9	260	25	ABX27892
32	81	2.9	1925	24	ABL41874
33	81	2.9	2056	24	ABL41873
34	78.2	2.8	2089	24	ABK51974
35	78.2	2.8	2089	25	ABX93210
36	77.6	2.8	1482	24	ABZ12990
37	75.4	2.7	1766	24	ABL41872
38	75	2.7	1644	24	ABZ12515
39	75	2.7	1690	24	ABL41879
40	72.6	2.6	2134	21	AAC36954
41	65.8	2.4	1872	24	ABK51975
42	65.8	2.4	1872	25	ABX93211
43	65	2.3	2171	22	AAK94284
44	63.4	2.3	1914	24	ABK51971
45	63.4	2.3	1914	25	ABX93207

ALIGNMENTS

[Handwritten signature]

Soybean cDNA clone
cDNA encoding soyb
wheat cDNA clone w
cDNA encoding whea
Corn tassels-deri
Rice cDNA clone rl
cDNA encoding rice
Corn ear-derived p
wheat cDNA clone w
cDNA encoding whea
DNA sequence from
Rice contig encodi
cDNA encoding rice
Corn tassels-deri
Soybean contig enc
cDNA encoding soyb
Arabidopsis thalia
Arabidopsis thalia
Nucleotide sequenc
Rice contig encodi
cDNA encoding rice
Arabidopsis thalia
Human GDP-mannose
Nucleotide sequenc
Nucleotide sequenc
wheat cDNA clone w
cDNA encoding whea
Arabidopsis thalia
Nucleotide sequenc
Nucleotide sequenc
Arabidopsis thalia
Nucleotide sequenc
wheat cDNA clone w
cDNA encoding whea
Human full-length
Corn cDNA clone ce
cDNA encoding corn

RESULT 1	ABK51962	standard; cDNA; 2824 BP.
ID	ABK51962	
XX	ABK51962;	
AC	ABK51962;	
XX		
DT	27-AUG-2002	(first entry)
XX		
DE	Corn contig encoding Arabidopsis thaliana-like sugar transport protein.	
XX		
KW	Corn; Arabidopsis thaliana-like sugar transport protein;	
KW	carbohydrate transport; grain filling; annual field crop;	
KW	plant; gene; ss.	
XX		
OS	Zea mays.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	238..2481
FT		/*tag= a
FT		/product= "Corn Arabidopsis thaliana-like sugar transport protein"
XX		
PN	US6383776-B1.	
XX		
PD	07-MAY-2002.	
XX		
PF	14-APR-1999;	99US-0291922.
XX		
PR	24-APR-1998;	98US-083044P.
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E I.	

XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2002-451386/48.
DR P-PSDB; AAU97201.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution
XX
PS Claim 3; Column 25-28; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC representing a contig assembled from various corn cDNA clones encodes
CC an Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other;

Query Match 79.2%; Score 2200.2; DB 24; Length 2824;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;

QY 133 CTTCACAGAGCCCGGTGACCTCGACGATATCTTGAGACAAAGATGCGGGCTGTTTC 192
Db 194 CTGGCCCGCGCTGCGGAGTGGGCGGTGATTTCCGGCGCCATGGGGGCGCCGTGA 253
QY 193 TTGTCCCATAGTCGCTCCATCGCAATCTATTCAGGGGTGGGCAATGCCACCATCG 252
Db 254 TGTGGCCATCGCGGCTCTATCGCAATCTGCAAGGCTGGGCAATGCGACAATTG 313
QY 253 CAGCTGCTGTTCTGATATATAAAGAAATTCAAATTGCAAAATGAGCCCACTGTGAGG 312
Db 314 CTGAGACCGCTCTGATACATAAAGAAATTCAACTGACAGAGCGACTTGATCGAAG 373
QY 313 GACTAATTTGTCATATGTCATCTATCGCGCCCAACCATGTTACTACATTTCCGGGCCAT 372
Db 374 GCGTCATCGTCCCATGTTCTCATTTGGGGCAACAGTCATCACACATCTCCGGGGCCAA 433
QY 373 TATCAGACTCGATTGGCCGACGCCCTATGCTTATTTCTCTTCAATTTCTGTACTTTCA 432
Db 434 GGGCTGACTGCGTTGGTAGAGGCCCATGCTGCTCGGCTGCTCTACTTCTGCTCA 493
QY 433 GCGGCTCATCATGCTATGCTCTCTAATGTCATGTCCTGCTGTTGGCAGCGCTTGAG 492
Db 494 GTGGGCTGATGCTTTGGGCGCAATGTCATCTTGTCTCTCGCAAGGCTCATTG 553
QY 493 ATGGATTTGGTATTTGCTTGCTGACGCTTGTGCTTTGTACATTTCAGAATAGCCC 552
Db 554 ATGGGTCGGTATCGGTTTGCGGTACACTTGTCTCTCTACATCTCCGAACCTGCAC 613
QY 553 C--TTCCGAGATTAGAGGTTTGTGATACTACCAATTCAGTGG--ATCAGGAG 606
Db 614 CGACACAGANAATCTTGGGGCTGNTNGAACACGTTGCCGAGTTCAITGGGGCTCAGNGAG 673
QY 607 GAATGTTCTTGTCACTGTCATGCTGTTGGGATGTCCTGTCGCATCACCCGATTGA 666
Db 674 GGATGTTCTCTCTACTGATGAGTGTGGGATGTCCTCATGCCCAAACTGATTGA 733
QY 667 GAATTTATGCTGTGTGCTCGGATACCTTCAATGTTCTTCTTGTGTTGACAATATTT 726
Db 734 GGTCTATGCTTGAAGTTCTGTGATCCCGTCACTTATNTACTTGTGACTGACTGTCTCT 793
QY 727 ATCTTCTGATCTCCAAAGATGCTGTTAGCAAAAGTCGATGGCAGAGCAAAAAAG 786
Db 794 ACTTGCTGATACCAAGGTGCTGTGNAAGCAAAAGAGATGGCGGAGGCAAGAGAG 853

QY 787 TGTTCAAAAAGTTACGGGGGAAAGACGATGCTCAGGTGAATGTCCCTTCTTTCGAG 846
Db 854 TGNTCAAAAGGCTGCGGGGAAAGAAAGATGTCANGGAGANGGCTCTTCTAGTTGAAG 913
QY 847 GGTTCAGGTTGAGAGACACTTTCATTGAAGAGTACATCTTGGACCTGCCACGAGG 906
Db 914 GTTGGGGGTGCGTAAAGATACAGTATTNNAGATACATTTGACCTGCCACGAGG 973
QY 907 CAGCCGATGATCTTGTACTGACGGTGATGAAGAAATACACTTTATGGGCGTGAAG 966
Db 974 CAGCCGATGATCTTGTACTGACGGTGATGAAGAAATACACTTTATGGGCGTGAAG 1033
QY 967 AAGCCAGTCATGATTTGCTGACCTTCTAAGGAGCCCATCATGCTTGAAGTGTCTTT 1026
Db 1034 AAGCCAGTCATGATTTGCTGACCTTCTAAGGAGCCCATCATGCTTGAAGTGTCTTT 1093
QY 1027 CTCTTGCATCTGTCATGGGAGCATGTGAACCAAGTGTACCCCTTATGATCCGATTG 1086
Db 1094 CTCTTGCATCTGTCATGGGAGCATGTGAACCAAGTGTACCCCTTATGATCCGATTG 1153
QY 1087 TGACACTTTTGTGATGTCATGAGATATATGCTCAAGCTGGAGGAAGTATGAGAGCA 1146
Db 1154 TGACACTTTTGTGATGTCATGAGATATATGCTCAAGCTGGAGGAAGTATGAGAGCA 1213
QY 1147 CATGTTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGCATGCCAAAATGAGC 1206
Db 1214 CATGTTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGCATGCCAAAATGAGC 1273
QY 1207 AGTGGATGAAGAGATCTTCATAGGGATGACGAGAGTACCATCTGATGTGACGAG 1266
Db 1274 AGTGGATGAAGAGATCTTCATAGGGATGACGAGAGTACCATCTGATGTGACGAG 1333
QY 1267 GTGACTATGAGACAATCTTCATAGCCCATTTGCTGTCCAGGCAAGCAAGTGGGAA 1326
Db 1334 GTGACTATGAGACAATCTTCATAGCCCATTTGCTGTCCAGGCAAGCAAGTGGGAA 1393
QY 1327 GGAAGGACATTTGTGACCATGCTCACCGTGAAGTCTTTGAGCATGAGAAGCAACC 1386
Db 1394 GGAAGGACATTTGTGACCATGCTCACCGTGAAGTCTTTGAGCATGAGAAGCAACC 1453
QY 1387 TCTTAGGGAGGTTGAGATGCTGTGAGCAGCATGATATCGTGGGAGTGGACCTTG 1446
Db 1454 TCTTAGGGAGGTTGAGATGCTGTGAGCAGCATGATATCGTGGGAGTGGACCTTG 1513
QY 1447 CTGGAATGTTGAGAGGAAGGTGAGAATGTTGAAAGAGGTGTTCAAAAGAG 1506
Db 1514 CTGGAATGTTGAGAGGAAGGTGAGAATGTTGAAAGAGGTGTTCAAAAGAG 1573
QY 1507 TCTACTTGACCAAGAGGATCTCTGGCTCAAGAGGGCTCAATTGTTTCACTTCCCG 1566
Db 1574 TCTACTTGACCAAGAGGATCTCTGGCTCAAGAGGGCTCAATTGTTTCACTTCCCG 1633
QY 1567 GTGCTGGCGATGTTTGAAGGTTAGTGAAGTTGTACATGCTGCTTTAGTAACTCACT 1626
Db 1634 GTGCTGGCGATGTTTGAAGGTTAGTGAAGTTGTACATGCTGCTTTAGTAACTCACT 1693
QY 1627 CAGCACTTTCTCAAGGGTCTGCTGAACCAAGCATGTCAGATGCTGCCATGTTCAAC 1686
Db 1694 CAGCACTTTCTCAAGGGTCTGCTGAACCAAGCATGTCAGATGCTGCCATGTTCAAC 1753
QY 1687 CATCTGAGGTAGCTGCCAAAGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAAGC 1746
Db 1754 CATCTGAGGTAGCTGCCAAAGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAAGC 1813
QY 1747 GTGCTCTGTTAGTCCGCTGTGGAATTCAGATCTTCAACAGTTTGTGGAATTAACGTTG 1806
Db 1814 GTGCTCTGTTAGTCCGCTGTGGAATTCAGATCTTCAACAGTTTGTGGAATTAACGTTG 1873
QY 1807 TTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGTGCGAGTTATTTTCCAAAT 1866
Db 1874 TTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGTGCGAGTTATTTTCCAAAT 1933

QY 1867 TTGCTCTACGCTCGGCATCAGCATCCATCTTGATCAGTCTCTCTCACTACTACTAATGC 1926
XX |||||
Db 1934 TTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTCTCTCTCACTACTAATGC 1993
QY 1927 TTCCTTGCAATTGGCTTTGCCATGCTGCTTATGGATCTTCCGGAAGAGTTTGTGCTGC 1986
XX |||||
Db 1994 TTCCTTGCAATTGGCTTTGCCATGCTGCTTATGGATCTTCCGGAAGAGTTTGTGCTGC 2053
QY 1987 TAGGCACAATTCCAATCTTGATAGCATCTAGTTATCCGTGTTGTCCTAATCTAATG 2046
XX |||||
Db 2054 TAGGCACAATTCCAATCTTGATAGCATCTAGTTATCCGTGTTGTCCTAATCTAATG 2113
QY 2047 ATTTGGGTACACTAGCCCATGCTTGTCTCCACCGCTCAGTGTATCGTCTACTTCTGCT 2106
XX |||||
Db 2114 ATTTGGGTACACTAGCCCATGCTTGTCTCTCCACCATCAGTGTATCGTCTACTTCTGCT 2173
QY 2107 GCTTCGTATGGGATTTGGTCCCATCCCGACATTTTATGTGCAAGATCTTCCAAACCA 2166
XX |||||
Db 2174 GCTTCGTATGGGATTTGGTCCCATCCCGACATTTTATGTGCAAGATCTTCCAAACCA 2233
QY 2167 GGGTTCGTGCTCTCTGTATTTGCCATTTGCTCTTACATTCGTGATCGGAGATATCATCG 2226
XX |||||
Db 2234 GGGTTCGTGCTCTCTGTATTTGCCATTTGCTCTTACATTCGTGATCGGAGATATCATCG 2293
QY 2227 TCACCTACAGCCTTCTCTGTGATGCTGAATGCTATTGGACTGGCGGCTTTTTCAGCATAT 2286
XX |||||
Db 2294 TCACCTACAGCCTTCTCTGTGATGCTGAATGCTATTGGACTGGCGGCTTTTTCAGCATAT 2353
QY 2287 ATGCAGTCGTATGCTTGATTTCTTTGTTGCTCTCTCTTAAGTCCCTGAGACAAAG 2346
XX |||||
Db 2354 ATGCAGTCGTATGCTTGATTTCTTTGTTGCTCTCTCTTAAGTCCCTGAGACAAAG 2413
QY 2347 GGATGCCCCCTTGAGGTATATTACCGAATCTTTGCAGTGTGTCGGAAGCAAGCGGCTGCA 2406
XX |||||
Db 2414 GGATGCCCCCTTGAGGTATATTACCGAATCTTTGCAGTGTGTCGGAAGCAAGCGGCTGCA 2473
QY 2407 AAGCCTAATTTCTTTGGTACCTTTGTGCAACTATTGCACGTGAAGTTAGAACTTGAA 2466
XX |||||
Db 2474 AAGCCTAATTTCTTTGGTACCTTTGTGCAACTATTGCACGTGAAGTTAGAACTTGAA 2533
QY 2467 GGGTTTCACCAAGAAGCTCGAGAAATTACTTTGATTGTGTAATGTTAAGGAACGA 2526
XX |||||
Db 2534 GGGTTTCACCAAGAAGCTCGAGAAATTACTTTGATTGTGTAATGTTAAGGAACGA 2593
QY 2527 ACATCTGCTCATGCTCTCTCAACCGGTAAAGAGTCCCTCAATGCGAATAGGAGTCGT 2586
XX |||||
Db 2594 ACATCTGCTCATGCTCTCTCAACCGGTAAAGAGTCCCTCAATGCGAATAGGAGTCGT 2653
QY 2587 TAAGTTGTCAATGTCATTACCAATATGTTTACCTATTGTACTGTATTATAAGTCAAGC 2646
XX |||||
Db 2654 TAAGTTGTCAATGTCATTACCAATATGTTTACCTATTGTACTGTATTATAAGTCAAGC 2713
QY 2647 TATTCAACGCTGCTGTTGTGCTAGAAATCTTAGAACAAAGAGATGAATGATCTGATCTGA 2706
XX |||||
Db 2714 TATTCAACGCTGCTGTTGTGCTAGAAATCTTAGAACAAAGAGATGAATGATCTGATCTGA 2773
QY 2707 TGTATATAATTAATCAAAATCTCAAAATAAGAAATATCGTTTCTCAAAAAAAA 2757
XX |||||
Db 2774 TGTATATAATTAATCAAAATCTCAAAATAAGAAATATCGTTTCTCAAAAAAAA 2824
RESULT 2
ABX93198 standard; cDNA; 2824 BP.
XX AC ABX93198;
XX AC
DT 29-MAY-2003 (first entry)
XX DE cDNA encoding corn sugar transport protein #1.
XX KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;

KW carbohydrate distribution; plant; gene; ss.
XX
OS Zea mays.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALLEN/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINNEY/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08326.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein
XX
PS Claim 2; Page 13-15; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other;
Query Match 79.2%; Score 2200.2; DB 25; Length 2824;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;
QY 133 CTTACAGAGAGCCGGTGACCTCGAGCATATCTTGAGAGCAAGATGTGGGGCTGTC 192
XX |||||
Db 194 CTTGGCCGCGCTGCGGAGTGGGGCGGTAGATTCCGGCGGCATGGGGCGCGTGA 253
QY 193 TTGTCCCATAGTCGCTCCATCGGCAATCTATTGCAAGGGGTGGACAATGCCACATCG 252
XX |||||
Db 254 TTGTCCCATAGTCGCGCTCTATCGGCAACTTGTGCAAGGGCTGGACAATGCCAATG 313
QY 253 CAGCTGCTGTTCTGTATATAAGAGAAATTTCAATTGCAAAATGAGCCCATGTGAGG 312
XX |||||
Db 314 CTGAGCCGCTCTGTATATAAGAGAAATTTCAACTGCAAGCGAGCCTGTGATGAAG 373
QY 313 GACTAATTTGTCAATGTCATTAATCGGCGCCACCATGCTTAATCTCTGATCTCTTCA 372
XX |||||
Db 374 GCTTCATCGTCCCATGTTCTCATTTGGGGCAACAGTCAACAACATCTCCGGGGCCAA 433
QY 373 TATCAGACTGATTTGGCCGAGCCCTATGCTTATTTCTCTCAATCTGTACTTCTTCA 432
XX |||||
Db 434 GGGTGAAGTGGTGGAGAGGCCCATGCTGCTGCGCTCGGCTGCTTACTTCTGCTCA 493
QY 433 GCGGCTCATCATGCTATGCTCTCTAATGTCATGCTGCTGCTGTCGACGCTTCGTAG 492
XX |||||
Db 494 GTGGGCTGTGATGCTTTGGGCGCCCAATGTGTACATCTTGTCTCTCGCAAGGCTCAT 553
QY 493 ATGATTTGTGATTTGGCTTGGCTGTACGCTTGTGCTTTGTACATTTCAAGAAATAGCCC 552

Db 554 ATGGGTTCGATCGGTTTGCGGTACACTGTTCCTCTACATCTCCGAACTGCAC 613
Qy 553 C--TTCCGAGATTAGAGGTTTGCTGAATACCTACCAATTCAGTG--ATCAGAG 606
Db 614 CGCACAGANATCTTGGGGCTGNTNGAACGTTGCCGAGTTCATTGGGGTTCAGNGAG 673
Qy 607 GAATGTTCTGTCTACTGTCATGCTGTTGGGATGTCCTGTGCCATCACCCGATTGGA 666
Db 674 GGATGTTCCCTCTCTACTGATGCTGTTGGGATGTCCTCATGCCCAACCTGATTGGA 733
Qy 667 GAATTATGCTTGGTGTGCTCGCGATACCTTCATTGTCTCTTGGTTGACATATTTT 726
Db 734 GGCTCATGCTTGAGATTCTGTCGATCCCGTCACTTATNTACTTTGACTGACTGTCTCT 793
Qy 727 ATCTTCCTGAATCTCCAAGATGCTCGTTAGCAAAAGTCGGATGGCAGAGCAAAAAG 786
Db 794 ACTTGCTGAATCACCAAGGTGGCTTGTAAGCAAAAGAGATGGCGAGCGAAGAGAG 853
Qy 787 TGTTCGAAAAGTTACGGGGGAAAGACGATGTCTCAGTGAATGTCCTTCTTCGAAG 846
Db 854 TGNTGCAAGGCTGCGGGGAAGAGATGTCTCANGGAGANGCTCTTCTAGTTGAAG 913
Qy 847 GGTTCGAGTTGAGAGAGACACTTCATTGAAGATACATCATGSACTGCACTGCACCGAG 906
Db 914 GTTGGGGGTGCGGTAAGATACACGTATTTNAGATACATCATGACCTGCACCGAGG 973
Qy 907 CAGCCGATGATCTTGTACTGACGGTGAATAAGAACAAATCACACTTTATGGGCTGAAG 966
Db 974 CAGCCGATGATCTTGTACTGACGGTGAATAAGAACAAATCACACTTTATGGGCTGAAG 1033
Qy 967 AAGGCCAGTCATGGAATTGCTCGACCTTCTAAGGGACCCCATCATGCTTGGAAAGTGCTTT 1026
Db 1034 AAGGCCAGTCATGGAATTGCTCGACCTTCTAAGGGACCCCATCATGCTTGGAAAGTGCTTT 1093
Qy 1027 CTCTTGCACTCTGTCATGAGGACATGTTGAACCAAGATGTACCCCTTATGATCCGATTG 1086
Db 1094 CTCTTGCACTCTGTCATGAGGACATGTTGAACCAAGATGTACCCCTTATGATCCGATTG 1153
Qy 1087 TGACACTTTTGGTAGTGTCCATGAGAAATATGCTCAAGCTGAGAGATATGAGAGCA 1146
Db 1154 TGACACTTTTGGTAGTGTCCATGAGAAATATGCTCAAGCTGAGAGATATGAGAGCA 1213
Qy 1147 CATGTTCCAAACCTTTGGAAGTATGTCAGTGTACAGATCAGCATGCCAAAAATGAGC 1206
Db 1214 CATGTTCCAAACCTTTGGAAGTATGTCAGTGTACAGATCAGCATGCCAAAAATGAGC 1273
Qy 1207 AGTGGATGAAGAGAAATCTTCATAGGGAATGACGAGAGTACCGCATCTGATGTCAGAG 1266
Db 1274 AGTGGATGAAGAGAAATCTTCATAGGGAATGACGAGAGTACCGCATCTGATGTCAGAG 1333
Qy 1267 GTGACTATGAGGACAAATCTCCATAGCCCATGCTGTCCAGGCAAGCAAGTGGGAAG 1326
Db 1334 GTGACTATGAGGACAAATCTCCATAGCCCATGCTGTCCAGGCAAGCAAGTGGGAAG 1393
Qy 1327 GGAAGCATGTGTCACCATGTCACCGTGAAGTGTCTTGAGCATGAGAAGGCAAAACC 1386
Db 1394 GGAAGCATGTGTCACCATGTCACCGTGAAGTGTCTTGAGCATGAGAAGGCAAAACC 1453
Qy 1387 TCTTAGGGGAGGTGAGATGCTGTAGCAGCACTGATATCGGTGGGGGATGGCAGCTTG 1446
Db 1454 TCTTAGGGGAGGTGAGATGCTGTAGCAGCACTGATATCGGTGGGGGATGGCAGCTTG 1513
Qy 1447 CTTGAAATGCTCAGAGAAGAGTGAATGTTAGAAAGAGAGTGTTCAAAAGAG 1506
Db 1514 CTTGAAATGCTCAGAGAAGAGTGAATGTTAGAAAGAGAGTGTTCAAAAGAG 1573
Qy 1507 TCTACTTGCAACAAGAGGAGTTCCTGGCTCAAGAAGGGGCTCAATTGTTCACTTCCCG 1566
Db 1574 TCTACTTGCAACAAGAGGAGTTCCTGGCTCAAGAAGGGGCTCAATTGTTCACTTCCCG 1633
Qy 1567 GTGTGGCGATGTTTTTGAGGGTAGTGATTGTACATGCTGCTTTAGTAAGTCAGT 1626

Db 1634 GTGTGGCGATGTTCTTGAGGGTAGTGATTGTGTACATGCTGCTTTAGTAAGTCAGT 1693
Qy 1627 CAGCACTTTTCTCAAAGGGCTTGCTGGAACCAAGCATGTCAAGATGCTGCCATGGTTCAAC 1686
Db 1694 CAGCACTTTTCTCAAAGGGCTTGCTGGAACCAAGCATGTCAAGATGCTGCCATGGTTCAAC 1753
Qy 1687 CATCTGAGGTAGCTGCCAAAGGTTACAGTTGGAAGATTTGTTGAACCTGGAGTGAAGC 1746
Db 1754 CATCTGAGGTAGCTGCCAAAGGTTACAGTTGGAAGATTTGTTGAACCTGGAGTGAAGC 1813
Qy 1747 GTGCCCTGTAGTCGGTGTGGAATTACAGATCTTCAACAGTTTGCTGGAATAAAGGTG 1806
Db 1814 GTGCCCTGTAGTCGGTGTGGAATTACAGATCTTCAACAGTTTGCTGGAATAAAGGTG 1873
Qy 1807 TTCTGTACTATACCCACAATTTCTTGAGCAAGCTGCTGCGCAGTTATCTTCCAAAT 1866
Db 1874 TTCTGTACTATACCCACAATTTCTTGAGCAAGCTGCTGCGCAGTTATCTTCCAAAT 1933
Qy 1867 TTGCTCTAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGC 1926
Db 1934 TTGCTCTAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGC 1993
Qy 1927 TTCTTGCAATTGGCTTTGGCATGCTGCTATGATCTTTCCGGAAGAGTTTTGCTGC 1986
Db 1994 TTCTTGCAATTGGCTTTGGCATGCTGCTATGATCTTTCCGGAAGAGTTTTGCTGC 2053
Qy 1987 TAGGCACAATTCGAATCTGATAGCATCTCTAGTTATCTGCTGTGTGCCAATCTAATTG 2046
Db 2054 TAGGCACAATTCGAATCTGATAGCATCTCTAGTTATCTGCTGTGTGCCAATCTAATTG 2113
Qy 2047 AATTGGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATCGTCTACTTGTCT 2106
Db 2114 AATTGGGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATCGTCTACTTGTCT 2173
Qy 2107 GCTTCGTATGGAATTTGGTCCCATCCCAACATTTATGTGCAGAGATCTTCCAAACA 2166
Db 2174 GCTTCGTATGGAATTTGGTCCCATCCCAACATTTATGTGCAGAGATCTTCCAAACA 2233
Qy 2167 GGGTTCGTGGCCTCTGTATGGCCATTTGTGCTTTACATTTCTGGATCGAGATATCATCG 2226
Db 2234 GGGTTCGTGGCCTCTGTATGGCCATTTGTGCTTTACATTTCTGGATCGAGATATCATCG 2293
Qy 2227 TCACCTACAGCCTTCGTGATGCTGAATGCTATTTGACTGCGGGGTGTTTCAGCATAT 2286
Db 2294 TCACCTACAGCCTTCGTGATGCTGAATGCTATTTGACTGCGGGGTGTTTCAGCATAT 2353
Qy 2287 ATGCAGTCGATGCTGATTTCTTGTGCTTCTTAAGTTCCTGAGCAAAAG 2346
Db 2354 ATGCAGTCGATGCTGATTTCTTGTGCTTCTTAAGTTCCTGAGCAAAAG 2413
Qy 2347 GGAATGCCCTTGAGGTTAATACCGAATCTTTGAGTGTGCGAAGCAAGCGGCTGCAA 2406
Db 2414 GGAATGCCCTTGAGGTTAATACCGAATCTTTGAGTGTGCGAAGCAAGCGGCTGCAA 2473
Qy 2407 AAGCCTAATTTCTTGGTACTTGTGTGCAACTATTGCACTGTAAGTTAGAACTTGAA 2466
Db 2474 AAGCCTAATTTCTTGGTACTTGTGTGCAACTATTGCACTGTAAGTTAGAACTTGAA 2533
Qy 2467 GGGGTTTCAACGAAGCTCGGAGAAATTACTTTGGAATTTGTGTAATGTTAAGGGAACGA 2526
Db 2534 GGGGTTTCAACGAAGCTCGGAGAAATTACTTTGGAATTTGTGTAATGTTAAGGGAACGA 2593
Qy 2527 ACATCTGCTCATGCTCTCAAAAGGTAAAGAGAGTCCCTCAATGGCAATAGAGTCTGT 2586
Db 2594 ACATCTGCTCATGCTCTCAAAAGGTAAAGAGAGTCCCTCAATGGCAATAGAGTCTGT 2653
Qy 2587 TAAGTTGTCATATGCTATTACCATATGTTTACCTTATGTAATTAAGTCAAGC 2646
Db 2654 TAAGTTGTCATATGCTATTACCATATGTTTACCTTATGTAATTAAGTCAAGC 2713
Qy 2647 TATTCAACGCTGTTGTTGCTAGAATCTTTAGAACAAAGATGATAATGATCTGATCTGA 2706
Db 2714 TATTCAACGCTGTTGTTGCTAGAATCTTTAGAACAAAGATGATAATGATCTGATCTGA 2773


```
QY 1371 ATGAGAGGCA--AACCTCTTAAAGGAGGGGTGNGATGGTGTGAGCAGCACTGATATC 1427
    |||||
Db 1357 ATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAAGTGAGCAAGGTGTAGTACAGTATT 1416
QY 1428 GGTGGGGGATGGCAGCTTGCTTGAATGTGTGAGAGAAGGAGTGAATGTAGAAAG 1487
    |||||
Db 1417 GGTGGGCTGGCAACTGCATGGAATGGACTGATAA--AGGTGAGATGAAAAACA 1473
QY 1488 GAAGTGGTTTCAAAGAGTCTACTTGCAACCAAGAGGAGTCTCGCTCAAGAGGGGC 1547
    |||||
Db 1474 CAAGGAGGGTTTAAAGAGATTATTAATGAGAGAGGAGTTCTGCATCTCGTGTGA 1533
QY 1548 TCAATTGTTTCACTTCCCGGTGGTGGCGATGTTTGTGAGGGTAGTGAAGTTGTACATGCT 1607
    |||||
Db 1534 TCCATTGTATCGATTCCCGGTGAAGCGGA-----TTGTCCAGGCT 1575
QY 1608 GCTGCTTTAGTAAGTCAAGTCAAGCACTTTTCTCAAAGGGTCTTGCTGAACCAAGCATGTCA 1667
    |||||
Db 1576 GCTGCTTGTGTAAGCCAAACCGCTCTTTACTCCAAAGAGCTTATGTGACACCAAGCTT 1635
QY 1668 GATGCTGCCATGGTTCAACCCATCTGAGGTAGCTGCCAAAGGTTCAGTTGGAAGATTTG 1727
    |||||
Db 1636 GGGCTTCAATGTTTCAACCCATCTGAGACAGCTTCAAGGGGCCAAGTTGGAAGCTCTT 1695
QY 1728 TTTGAACCTGAGTGAAGGCGTCCCTGTAGTCGGTGTGAATTCAGATCCTTCAACAG 1787
    |||||
Db 1696 CTGGAACCAAGGGTTAAGCATGCAATTGGTGTGTGAGTTGAATACAATACTTCAAGCAG 1755
QY 1788 TTTGCTGAATAAACGGTGTCTGTACTATATACCCCAACAAATCTTGAGCAAGCTGTGTG 1847
    |||||
Db 1756 TTTTCAGGGATAAATGGGGTCTATATTAACACACCTCAAACTCTGAAGAGGCCGGTGT 1815
QY 1848 GCAGTTATTTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCT 1907
    |||||
Db 1816 GAGTTCTTTCTTCAAGATATAGGCAATTGGCTCAGAGTGGCATCATCTTATCAGTGTCT 1875
QY 1908 CTGACTACCTTACTAATGCTTCTTGCAATGGCTTGGCCATGCTGCTTATGATCTTCC 1967
    |||||
Db 1876 TTGACAACTTCTTGATGCTTCTGTATAGGCGTAGGCCATGAAGCTCATGAGTGTCA 1935
QY 1968 GGAAGAAGGTTTTGTGCTAGGCACAATTCGAATTTGATAGCATCTCTAGTTATCTG 2027
    |||||
Db 1936 GGCAGAAAGCAGTGTCTACTACTAACAATCCCGTGTGATGTGTCACTCATATTATTG 1995
QY 2028 GTTGTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCACCGTCACT 2087
    |||||
Db 1996 GTGATTTGGAAGCCTGTGAATTTTGGCAATGTGCGCCATGAGCAATCTCAACAGTATGC 2055
QY 2088 GTTATCGTCTACTTCTGCTGCTTGTATGGGATTTGGTCCCATCCCAACATTTATGT 2147
    |||||
Db 2056 GTTGTGTTTATTTCTGCTGCTTGTGATGGGTATGACCAATTTCCAAACATCTTGTGC 2115
QY 2148 GCAGAGATCTTCCAAACCAAGGTTCTGCGCTCTGTATGGCATTTGTGCTTATCATTC 2207
    |||||
Db 2116 TCAGAGATTTTCCCACTAGGGTGGCTGCGCTGTGATGCTATCTGTGATAGTGTTC 2175
QY 2208 TGGATCGAGATATCATGCTCACTACAGCCTTCTGTGATGCTGAATGCTATTTGACTG 2267
    |||||
Db 2176 TGGATTTGAGACATCATCATCACTACTGCTGCTGTGATGCTCGGCTCTTTAGGACTT 2235
QY 2268 GCGGTTGTTTCAAGCATATATGAGTGTATGCTGATTTCTTGTGTTGCTTCTCTT 2327
    |||||
Db 2236 GGTGGTGTATTTGCCATTTAGCGAGTTGTGTTTCACTCGTGATATTTGTGTTTTG 2295
QY 2328 AAGTCCCTGAGACAAAGGGGATGCCCCCTGAGGTTATTAACCAATTTCTTGACAGTTGCT 2387
    |||||
Db 2296 AAGTTCAGAAACAAAGGGGATGCCCCCTTGAAGTCACTCTGAATTTCTTGTGTGA 2355
QY 2388 GCGAAGCAAGCGGCTGC 2404
    |||||
Db 2356 GCAAAGCAGGCTGCTTC 2372
    |||||
```

```
RESULT 4
ABX93201
ID ABX93201 standard; cDNA; 2601 BP.
XX
AC ABX93201;
XX
DT 29-MAY-2003 (first entry)
XX
DE cDNA encoding soybean sugar transport protein #1.
XX
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Glycine max.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08329.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
XX
PS Claim 2; Page 19-20; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other;

Query Match 33.9%; Score 940.2; DB 25; Length 2601;
Best Local Similarity 66.0%; Pred. NO. 1.6e-212;
Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

QY 177 ATGTGGGGGCTTCTGTGCGCATAGTGCCTCCATCGGCAATCTATGAGGGGTG 236
    |||||
Db 175 ATGAAGGTGCGGCTCTGTGCTATTTGCCGCTCCATTTGTAATTTCTTCCAAGATGG 234
QY 237 GACAATGCCACCATCGCAGCTGCTGTCTGTATATAAAGAAGGAATTTCAATTGCCAAAT 296
    |||||
Db 235 GATAATGCTACCATCGCCGGGCTAATGTTAATTAAGAAAGACCTTGTGCGAA-- 292
QY 297 GAGCCCACTGTGGAGGGAATAATGTGTCAATGCTATATCGGCGCCACCATCTTACT 356
    |||||
Db 293 ----CACTATGAAAGGCTTGTGTGGGCAATGCTCCGATTGGAGCAACGGTAATCACC 348
    |||||
```

OY		357	ACATTCCTCCGGGCCATTATCAGACTCGATTGCGCCGACGCCCTATGCTTATTCCTCTCA	416
Db		349	ACATGCTCTGTCTCATAGCGGAATTGGCTCGGTCGCACCCATGATGATAATCTCACT	408
OY		417	ATTCTGACTTCTTGAGCGGCTCATCATGCTATGGTCTCCTAATGCTATGTCCTG	476
Db		409	GTCCTCATTTCTTGGGTGGTTGGTGATGCTGTGGTCCCCAAATGTGATGTGTGTC	468
OY		477	TTCGCACGCTTCGTAGATGAAATTTGGTATTGGCTTGGCTGTCA CGCTGTGCTTTGAC	536
Db		469	TTGGCGAGGCTACTTGATGGAATTTGGGATTGGCCCTTGCTGTGACTCTGTGCCGCTAT	528
OY		537	ATTTCAGAAATAGCCCCCTTCGGAGATTAGAGTTTGCTGAATACACTACCACAATTGAGT	596
Db		529	ATATCTGAAAACGGCGCCGCTCTGAATAAGGGGTCGTGAATAACGCTTCCTCAGTTCACT	588
OY		597	GGATCAGAGAGAATGTTCTGT CATACTGCATGGTGTGGGATGTCCTGTCCCACTCA	656
Db		589	GGCTCTGAGGAATGTTTTGTCTACTGTATGTTTGGCATGTCAATTGAGTCCC GCG	648
OY		657	CCCCATTGGAGATTAATGCTTGGTGTGCTCGCGATPACTTCAATTGTTCTTCTTGGTTTG	716
Db		649	CCTAGCTGAGGCTCATGCTTGGGGTCTGTCTATTTCTTCTCTGTATTTTGCA TTG	708
OY		717	ACAATATTTTATCTTCTGAACTCCCAAGATGGCTCGTTAGCAAAGGTCGAGTGCAGAG	776
Db		709	ACCATT TTTTCTTGCCCGAGTCTCTCGGTGCTGTCA GCAAGAGATGCTCGAG	768
OY		777	GCAAAAAAGGTGTGCAAAAGTTACGGGGAAAGACGATGTT CAGGTGAATGTCCCTT	836
Db		769	GCTAAGAGGTGCTCCAAGATTGCGCGAAGGAGAGATGTGT CAGGCGAGATGGCA TTG	828
OY		837	CTTCTCGAAGGTTGAGGTTGAGAGACACTTCCATTGAAGAGTACATCATTTGACCT	896
Db		829	CTGTTGAAGTCTCGGATTTGGGGGTGATACATCTATCGAAGAGTACATAATTGG CCT	888
OY		897	GCCACCGAGGCGCCGATGATCTGTTACTGACGGTGAAGAA CAATCACACTTTAT	956
Db		889	GCTGACGATGTGCTGAATGTCATGAACATGCAACAGAAAGATAAAATTCGATTATAT	948
OY		957	GGGCTGAAGAGGCGCACTCATGATTGCTCGACCTTCTAAGGAGCCCATCATGCTTGA	1016
Db		949	GGATCCCAAGCAGGCTTCTTGTTATCAAAACCCTGTCACTGACAGAGTTCTATTTG-	1007
OY		1017	AGTGTCTTCTCTTGCACTCTCGTCATGGGAGCATGGTGA ACCAGCTGATCCCTTATG	1076
Db		1008	------CCTTGCCTCACACCATGGAGCATCATCAACCAAGCATGCCCTCATG	1056
OY		1077	GATCCGATTGTGACACTTTTGTAGTGTCCATGAGAAATATGCTCA-----AGCTGA	1130
Db		1057	GATCCTCTGTGACACTGTTTGTAGCATTCATGAGAAGCTCCCGAGACAGAGACA GA	1116
OY		1131	GGAAGTATGAGAGCACATGTTTCCAAACTTTGAAGTATGTT CAGTGTACAGATCAG	1190
Db		1117	GGAAGCATGCGAAGCACTGTGTTCCAAATTTTGAAGCATGTTCA GCACTGTGAGCCG	1176
OY		1191	CATGCCAAAATGACGAGTGGGATGAAGAGATCTTCA TAGGGATGACGAGAGTACGCA	1250
Db		1177	CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTHACATG	1236
OY		1251	TCTGATGTGACAGAGGTGACTATAGGACAATCTCCATAGCCCATGTGCTGTCCAGGAG	1310
Db		1237	TCAGATGCAACCCGTGGGACTCCGATGATTAATTGCAAGTCTT TAATCTCAGCCAA	1296
OY		1311	GCAACAGGTGCGAAGGGAAGACATTTGTGACACCATGTCA CCGTGAAGTCTTTGAGC	1370
Db		1297	ACAACAAGCCTTGAAAAAGACTTACCTCTCTCTCTCCATGGCAGATATCCTTGGCAGC	1356
OY		1371	ATGAGAAGCA--AACCTCTTAGGGGAGG GTGAGATGTTGTGACAGCACTGATATC	1427
Db		1357	ATGAGGCGTCAAGTATGTTCA TGCAAGGGTCAAGGTGAGCAAGGTGTAGTACAGGTA TT	1416
OY		1428	GGTGGGATGGCAGCTTGTCTTGA AATGTCAGAGAAAGGAGTGAATGTGTAGAAAG	1487

Db	1417	GGTGGCTGGCACTGGCATGGAAATGAGTGAATAA--AGGTGAGGATGGAAAAACAA	1473
QY	1488	GAAAGTGGTTTCAAAAAGTCTACTTGCACCAAGAGGAGTTCCGGCTCAAGAAAGGGC	1547
Db	1474	CAAGAGGGTTTAAAGAAATTTATTACATGAGAGGAGATTCTGCATCTCGTGTGA	1533
QY	1548	TCAATTGTTTCACTTCCCGGTGGTGGCGATGTTTTGAGGGTAGTAGTTGTACATGCT	1607
Db	1534	TCCATTGTATCGATTCCCGGTGAAGCGAA-----TTGTCCAGGCT	1575
QY	1608	GCTGCTTTAGTAAGTCAGTCAGCACTTTTCTCAAGGCTTGTCTGAACCAAGCATGTCA	1667
Db	1576	GCTGCCCTTGGTAAGCCAACCCGCTCTTACTCCAAAGAGCTTATTGATGGACACCCAGTT	1635
QY	1668	GATGCTGCCATGGTTCACCCATCTGAGGTAGCTGCCAAAGGTTCAACGTTGGAAGATTG	1727
Db	1636	GGGCTGCAATGGTTCACCCATCTGAGACAGCTTCAAAAGGGCCAAAGTTGAAAGCTCTT	1695
QY	1728	TTTGAACCTGGAGTGAAGCGCGTCCCTGTAGTCGGTGTGAATTAGATCTTCAACAG	1787
Db	1696	CTTGAACCAAGGGTTAAGCATGCATTGTTGTGAGTTGAATACAAATPACTTCAGCAG	1755
QY	1788	TTTGCTGAATAAAGCGTGTCTGTACTATACCACAAATCTTGAGCAAGCTGGTGTG	1847
Db	1756	TTTTCAAGGATAAATGGGGTCTATATTACACACCTCAATCCTTGAAGAGCGCGGTGT	1815
QY	1848	GCAGTTATTTCTTCCAATTTGCTCAGCTCGGCATCAGCATCCATCTTGATPAGTTCT	1907
Db	1816	GAAATTCTTTCTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATTTCTTATCAGTCT	1875
QY	1908	CTCACTACCTTACTTATGCTTCTCTGCATTGGCTTTGGCCATGCTGCTTATGATCTTCC	1967
Db	1876	TTCAACAACCTTCTGATGCTTCCCTGTATAGCGGTAGCCATGAAAGCTCATGATGTTTCA	1935
QY	1968	GGAAGAAGGTTTGTGCTGCTAGGACACAATTCAAATCTTGATAGCATCTTGTATTCCTG	2027
Db	1936	GGCAGAAGGCAGTTGCTACTTACTACAAATCCCGTGTGATTTGTCTCACTCATATTTTG	1995
QY	2028	GTTGTGTCCAATCTAATTGATTTGGGTACACATGAGCCCATGCTTGTCTCCACCGTCAGT	2087
Db	1996	GTCAATTGGAAGCCTGGTAATAATTTGGCAATGTGCGCCATGCAGCAATCTCAACAGTATGC	2055
QY	2088	GTTATCGTCTACTTCTGCTGCTTCTGTTATGGGAATTTGCTCCCATCCCCACAATTTATGT	2147
Db	2056	GTTGTGTTTATTTCTGCTGCTTGTGATGGGTTATGGAACCAATTCCAAACATCTTTGCTC	2115
QY	2148	GCAGAGATCTTTCCAACCAAGGTTGCTGGCCTCTGTATTGCCATTGTGTGCTTTACATTC	2207
Db	2116	TCAGAGATTTTCCCACTAAGGGTGCCTGCGCTCTGCATTGCTATCTGTGCATTAAGTGTTC	2175
QY	2208	TGGATCGAGATATCATCGTCAACCTACAGCCTTCTGTGATGCTGAATGCTATTGGACTG	2267
Db	2176	TGGATTGAGACATCATCATCACATACTCGCTGCTGTGATGCTCGGCTCTTAAAGACTT	2235
QY	2268	GCGGGTGTTTTCAGCATATATGCACTGATCGTATGATTTCCTTGTGTGCTTCTCCTT	2327
Db	2236	GGTGTGTATTCGCCATTAAAGCAGTGTGTTGTTTCATCTCGTGATATTGTGTTTTTG	2295
QY	2328	AAGTCCCTGAGACAAAGGGGATGCCCTTGAAGTTATTACCGAATCTTTGCAATTGCT	2387
Db	2296	AAGTTCAGAAACAAAGGGCATGCCCCCTTGAAGTCACTCTGAATTTCTTCTGTTGGA	2355
QY	2388	GCGAAGCAAGCGGCTGC 2404	
Db	2356	GCAAAGCAGGCTGCTTC 2372	

RESULT 5	
ABZ13510	
ID	ABZ13510 standard; DNA; 2190 BP.
XX	
AC	ABZ13510;

XX	21-JAN-2003	(first entry)	
DT	Arabidopsis thaliana stress regulated gene	SEQ ID NO 1315.	
XX	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.		
DE	Arabidopsis thaliana.		
XX	WO200216655-A2.		
XX	28-FEB-2002.		
XX	24-AUG-2001; 2001WO-US26685.		
XX	24-AUG-2000; 2000US-227866P.		
PR	26-JAN-2001; 2001US-264647P.		
PR	22-JUN-2001; 2001US-300111P.		
XX	(SCRI) SCRIPPS RES INST.		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX	Harper JF, Kreps J, Wang X, Zhu T;		
PI	WPI; 2002-304127/34.		
DR	Identifying a stress condition to which a plant cell has been exposed		
XX	and producing plants with increased tolerance to these abiotic stresses		
PT	-		
XX	Claim 144; SEQ ID NO 1315; 577pp + Sequence listing; English.		
PS	The invention relates to identifying a stress condition to which a plant		
XX	cell has been exposed, comprising:		
CC	(a) contacting nucleic acid representative of expressed polynucleotides		
CC	in the plant cell with an array or probes representative of the plant		
CC	cell genome; and		
CC	(b) detecting a profile of expressed polynucleotides in the plant cell		
CC	characteristic of a stress response. The method is useful in the		
CC	production of transgenic plants, cells and seeds and in producing plants		
CC	with increased tolerance to abiotic stress. The present sequence is that		
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used		
CC	in methods of the invention.		
CC	Note: The sequence data for this patent is not represented in the printed		
CC	specification but is based on sequence information supplied to Derwent by		
CC	the European Patent Office.		
XX			
SQ	Sequence 2190 BP; 529 A; 439 C; 572 G; 650 T; 0 other;		
XX			
XX	Query Match	31.5%; Score 875.8; DB 24; Length 2190;	
XX	Best Local Similarity	64.9%; Pred. No. 2.8e-197;	
XX	Matches 1449; Conservative	0; Mismatches 722; Indels 60; Gaps 8	
QY	177 ATGTCGGGGGCTGTCTTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGAGGGGTGG	236	
Db	1 ATGAGTGGAGCTGTGCTTGTGCTATGTGCTGCTGCTGCTGTGGCAACTGTGTACAAGATGG	60	
QY	237 GACCAATGCCACCATCGCAGCTGCTGTCTGTATATAAAGAAGAAATTTCATTGCCAAAT	296	
Db	61 GATTAACGCACTATTGCAAGAGCTGTGTGTACATAAAAAAGAGTTTAATTGGAGAGT	120	
QY	297 GAGCCCACTGTGAGGGGACTAATTGTGTCAATGTCACCTATTCGGGCCACCATCGTACT	356	
Db	121 AATTCATCAGTGAAGGTCTAATTGTGGCATGTCACCTATTGTGTCTCTGATTACA	180	
QY	357 ACATTTCTCGGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTTCA	416	
Db	181 ACATGCTCTGAGGGGTAGCTGATTTGGCTTGGTGGCCGCTCCCATGCTAATATGTCTCA	240	
QY	417 ATCTGTACTTCTTCAGCGGCTCATCATGCTATGTCTCTTAATGTCTATGTCTGCTG	476	
Db	241 ATCTCTACTTGTGTGCTTCTCTAGTAATGCTATGTCTCCGAATGTTATGTGTGCTC	300	

QY	477	TTGGCACCCTTCGTAGATGGATTGGTATTTGGCTTGGCTGTACGCTTGTGCTTTGTAC	536
Db	301	TTAGGAAGGTTGTTAGATGGAATTTGGGGTTGGTCTTGTGTGCACACTTGTCTTATTAT	360
QY	537	ATTTCAGAAATAGCCCCCTTCGGAGATTAGAGTTTGTCTGAATACACTACCAATTCAGT	596
Db	361	ATATCTGAGACTGCACCACCTGAGATTAGGGGACTGTGAATACGCTACCGCACTTCACT	420
QY	597	GGATCAGAGGAATGTTCTTGTCACTGATGATGTTTGGATGTCCCTGTCCCATCA	656
Db	421	GGCTCTGAGGAGTGTCTTATCTTACTGTATGTTTTCGGAATGTCTGTGATGCCATCA	480
QY	657	CCCGATTGAGAATTATGCTTGTGTGCTCGCATACTTCATTGTTCTTGTGTTG	716
Db	481	CCTAGCTGAGATTGATGTGTTGGTGTCTCTTTTCATCCCTTGTCTTTTCTTCCCTC	540
QY	717	ACAATATTTTATCTTCTCTGAATCTCCAAGTGGCTTCGTTAGCAAAGSTCCGATGCCAG	776
Db	541	ACGGTCTTCTTCTTCCCGAGTCCCAAGSTGGCTCTGTGACCAAAGSTCGAATGCTTGAA	600
QY	777	GCAAAAAAGGTGTTGCAAAAGTTACGGGGGAAAGCATGTCTCAGGTGAATGTCCCTT	836
Db	601	GCAAAAGCGGTTCTTTCAGAGACTGCGTGTGCGGAAGATGTCTGTGTAGATGGCTTTG	660
QY	837	CTTCTCGAAGGGTTGAGGTGGAGGAGACACTTCATTTGAAGATACATCATTTGACCT	896
Db	661	TTGCTTGAGGGTCTTGGAAATGGAGGTGAACCAACATAGAGAAATATTAATTTGGTCCC	720
QY	897	GCCACCGAGGACCCGATGATCTTGTACTGACGGTGAATAAGAACAAATCACACTTAT	956
Db	721	GCGGATGAAGTTACTGTATGATCATGATATAGCTGTGGAATAAGATCAATTAAGTTATAT	780
QY	957	GGGCTGGAAGAGGCCAGTCAATGGAATGCTCGACCTTTCTAAGGGACCCATCATGCTTGA	1016
Db	781	GGTGCAGAGAAGGGCTGAGTTGGGTTGCTAGGCCACTCAAAAGG-----GGA	828
QY	1017	AGTGTGCTTCTCTTGCATCTCGTCATGGGAGCATGATGTAACCAAG--TGTACCCCTT	1073
Db	829	AGCACTATGAGTGTGTTGTCTCGCCATGGAAGTACAAAGCAGGAGGCAAGGCTCATTG	888
QY	1074	ATGATCCGATTGTGACATTTTGTGTAAGTGTCCATGGAATATGCTCAAGCTGAGGA	1133
Db	889	ATTGATCCTCTTGTACACTGTTGGAGCGTTCAAGAAAGATGCCGACACT---GGA	945
QY	1134	AGTATGAGGAGCACATGTTTCCAACTTTGGAAGTATGTTCAAGTGTACAGATCAGCAT	1193
Db	946	AGCATGAGGAGTGCCCTGTCCACATTTTGGAGTAATGTTCAAGTTGAGGGAATCAA	1005
QY	1194	GCCAAATAGACAGTGGGATGAAGAAATCTTCATAGGGATGACGAGAGTACGCATCT	1253
Db	1006	CCAAGACATGAAGATTGGGATGAAGAAATCTTGTGAGAGAAGTGAGATTAATCCATCC	1065
QY	1254	GATGCTGACGAGGAGTACTATGAGCAATCTCCATAGCCCATTTGCTGTCCAGGCGCA	1313
Db	1066	GACCAT--GGAGATGATTTGAAAGATGATCTTCAATTTCCGTTGATCTCACGTCAAACG	1122
QY	1314	ACAGGTCCGGAAGGGAAGCAATTGTGCACCATGTGTACCCGTGGAAGTGTGAGCATG	1373
Db	1123	ACAAGCATGGA--GAAGACATGCTCTCACACTGCTCATGGAACCTTTTCTTACCTTCAGA	1179
QY	1374	AGAAGGCAAAACCTCTTAAAGGGAAGGTGAGATGTTGAGCAGCAGCACTGATATCGTGGG	1433
Db	1180	CATGGAAGTCAAGTGCAGGAGCTCAAGGGGAAGACGGGTAGTATGGGATTTGAGGT	1239
QY	1434	GGATGGCAGCTTGCTTGAATAGTTCAGAGAAGGAAGTGAATGTAGAAAGGAAGT	1493
Db	1240	GGATGGCAAGTGCATGGAATGGAACGAAAGAGAAGATGAATCCGGACAGAAAGAA	1299
QY	1494	GGTTTCAAAAAGTCTACTTGCAACCAAGAGGAGTTCCTGCTCAAGAAGGGCTCAATT	1553
Db	1300	GGTTTC-----CCAGGATCTCGACGTGGCTCAATT	1329
QY	1554	GTTTCACTTCCCGGTGTGGCGATGTTTGTGAGGGTAGTGAAGTTGTACATGCTGTGCT	1613


```

Db      1330 GTTCAATTGCTGTGTGATGAAACGGTGAGGCA--GATTTGTACAAAGCGTCTGCT 1386
QY      1614 TTAGTAAGTCAGTCAGCACTTTCTCAAGGGTCTTGCTGAACCAAGCATGTCAGATGCT 1673
Db      1387 TTGGTTAGCCCAACGAGCTCTTTATTCCAAAGACCTTCTCAAGAAACATACAAATTGGTCT 1446
QY      1674 GCCATGCTTCAACCCATCTGAGGTAGCTGCCAAAGGTTCAAGTTGAAAGATTGTTGAA 1733
Db      1447 GCTATGTAATCATCCATCCGAAACAACT--AAAGGGTCAATTGGCATGATCTTCATGAT 1503
QY      1734 CCTGAGTGAAGCGGCGTCCCTGTTAGTGGGTGTTGGAATTGAGATCTTCAACAGTTGCT 1793
Db      1504 CCTGAGTCAAGCGTGCAATGATGAGAGTTGAGCTTCAATTAATTGAGCATGTTCTCA 1563
QY      1794 GGAATAAACGGTGTCTGTACTATATACCCCAAAATTTCTGAGCAAGCTGTGTGAGTT 1853
Db      1564 GGCAATCAACGAGATTCTTACTACACACCGCAATCTTGAGCAGCGGGGTGTCGGGATC 1623
QY      1854 ATCTTTCCAATTTGTGTCTCAGCTGGGATCAGCATCCATCTTATGATCAGTCTCTACT 1913
Db      1624 CTACTATCGAACATGGGATAGTTCTTCTCAGCATCTTACTTATAAGTGCATTGACA 1683
QY      1914 ACCTTACTAATGCTTCTTGCATGCTGCTTGGCATGCTGCTGATGATCTTCCGAGAA 1973
Db      1684 ACCTTGTGATGTTTACCTGCAATAGCTGTGCAATGAGGCTCATGATCTTCTGTGCA 1743
QY      1974 AGGTTTGTGCTGCTAGGACAAATTCATCTGTATAGCATCTAGTTATCTGTTGTG 2033
Db      1744 AGGACCTGTCTTCTCACCACGATCAATCTGATAGCATCTGTTAGTTAATC 1803
QY      2034 TCCAAATCAATTTGATTTGGTACACAGCCCATGCTTGTGCTCCACCGTCAGTGTATC 2093
Db      1804 TCAATCTTGTTCACATGAACAGCATGTGTGACGCGGTCTTATCAACCGTAAGCTGTG 1863
QY      2094 GTCTACTTCTGCTGCTTCTGTTATGGGATTTGGTCCCATCCCAACATTTATGTGACAG 2153
Db      1864 CTCTACTTCTGCTTCTTCTGATGGGTTTCGGTCTGCTCCAAACATCTCTGTTCAGAG 1923
QY      2154 ATCTTTCCAACCAAGGTTCTGCTGCTCTGATTTGCCATTTGTCCTTATCATTTCTGATC 2213
Db      1924 ATTTTCCAACCTCGAGTCCGCGAATCTGCATCGCCATCTGCGCACTCACCTTCTGATC 1983
QY      2214 GGAGATATCATGCTCACCCTACAGCTTCTGCTGATGCTGAATGCTATTGAGCTGGCGGT 2273
Db      1984 TGTGACATTAATGCTCAGCTTACAGTCTCCCGTCTGCTCAATTCATTTGAGTACTGCT 2043
QY      2274 GTTTTACAGATATATGACAGTCTGATGCTGATTTCTTGTGTTGCTTCTTAAAGTTC 2333
Db      2044 GTGTTGGAATGTACGCAATCGTATGTTGATTTTCATGGGCTTGTGTTCAATAAGTTC 2103
QY      2334 CCTGAGCAAAAGGGGATGCCCTTGAGGTTATTACCGAATCTTTGCAAGTGTGCGAAG 2393
Db      2104 CCGGAATAAAGGCATGCCACTTGAAGTATCATCAGAGTCTTTTCTGTTGAGCTAGA 2163
QY      2394 CAAGCGGCTGC 2404
Db      2164 CAAGCTGAAGC 2174

```

RESULT 6

```

ABZ14449
ID      ABZ14449 standard, DNA; 2205 BP.
XX
AC      ABZ14449;
XX
DT      21-JAN-2003 (first entry)
XX
DE      Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.
XX
KW      Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS      Arabidopsis thaliana.

```

```

XX      XX
PN      WO200216655-A2.
XX      XX
PD      28-FEB-2002.
XX      XX
PF      24-AUG-2001; 2001WO-US26685.
XX      XX
PR      24-AUG-2000; 2000US-227866P.
PR      26-JAN-2001; 2001US-264647P.
PR      22-JUN-2001; 2001US-300111P.
XX      XX
PA      (SCRI) SCRIPPS RES INST.
PA      (SYGN) SYNGENTA PARTICIPATIONS AG.
XX      XX
PI      Harper JF, Kreps J, Wang X, Zhu T;
XX      XX
DR      WPI; 2002-304127/34.
XX      XX
PT      Identifying a stress condition to which a plant cell has been exposed
PT      and producing plants with increased tolerance to these abiotic stresses
PT
XX      XX
PS      Claim 144; SEQ ID NO 2254; 577bp + Sequence Listing; English.
XX      XX
CC      The invention relates to identifying a stress condition to which a plant
CC      cell has been exposed, comprising:
CC      (a) contacting nucleic acid representative of expressed polynucleotides
CC      in the plant cell with an array or probes representative of the plant
CC      cell genome; and
CC      (b) detecting a profile of expressed polynucleotides in the plant cell
CC      characteristic of a stress response. The method is useful in the
CC      production of transgenic plants, cells and seeds and in producing plants
CC      with increased tolerance to abiotic stresses. The present sequence is that
CC      of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC      in methods of the invention.
CC      Note: The sequence data for this patent is not represented in the printed
CC      specification but is based on sequence information supplied to Derwent by
CC      the European Patent Office.
XX      XX
SQ      Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 other;

```

```

Query Match      25.2%; Score 700; DB 24; Length 2205;
Best Local Similarity 61.2%; Pred. No. 1.2e-155;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;

```

```

QY      177 ATGCGGGGGTGTCTTGTGCGCATAGTCGCTCCATCGCAATCTATTGACGGGTG 236
Db      1 ATGAAGGAGGAGCACTCTGCTGCTCTCGCCGCAATCGGCAATTTCTTACAGAGATG 60
QY      237 GACATGCCACCATCGAGCTGCTGTCTGTATATAAGAAGATTTCATTTGCAAAAT 296
Db      61 GACATGCCACCATTTGCTGAGCTATGTTATATACAAAGACTTGAATCTACCAA-- 118
QY      297 GAGCCCACTGTGAGGAGCTAATTTGTCAATGTCACTTATCGGGCCACCATCGTTACT 356
Db      119 ----CCTCTGTCAAGGCTGTGCTGTGCTATGTCAATGATCGGTGCAACGGTCAATCA 174
QY      357 ACATTTCTCCGGCCATATCAGACTGATTGGCCGAGCCCTATGCTTATTTCTCTTCA 416
Db      175 ACTTGCTCAGGACCGATATGTATTTGCTCGGACAGCGCCCATGCTCATTTTATFATCA 234
QY      417 ATTCTGACTTCTTCAAGCGCTCATCATGCTATAGGTCTCTAATGTTATGTCCTGCTG 476
Db      235 GTTATGTATTTGCTGCGGTTGTATATGTTGTGCTCTCCCAATGTCTATGTTCTGTC 294
QY      477 TTGGCAGCTTCTGATAGGATTTGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
Db      295 TTGCTAGGCTTCTTATATGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
QY      537 ATTTCAGAAATAGCCCTTCCGAGATTGAGAGTTTGTGGAATACACTACCAATTCAGT 596
Db      355 ATTCTGAACCGCTCTCTCCGAGATCAGAGGACAGTTAAATCTCTCCCTCAGTTTCTT 414

```

OY		597	GGATCAGGAGGAATGTTCTTGTCATACTGCATGTGTGGATGTCCCTGTGCCATCA	656
Db		415	GCGTCTGGTGAATGTTTTTGTCATACTGTATGGTTTCACTATGTCCCTGAGTACTCC	474
OY		657	CCCCATTGGAGAATTANGCTTGGTGTGCTCGCGATACTTCATTGTTCTTTGGTTTG	716
Db		475	CCTAGCTGAGAGCCATGCTCGGTCTCTCGATCCCTTCTCTTATTATTGTTCTC	534
OY		717	ACAATATTTTATCTTCTGTAATCTCCAAGATGGCTCTTAGCAAAGTCGATGCGACAG	776
Db		535	ACGGTGTTTTATTTGCCGAGTCTCTCGTTGGCTGTAGTAAAGGAAGATGACAG	594
OY		777	GCAAAAAGGTGTTGCAAAAGTTACGGGGGAAGACSATGTCTCAGGTGAATGTCCCTT	836
Db		595	GCTAAGCGAGTCTTCTCAACAGTTATGTGGCAGAGAAAGATGTTAACCATGAGTGGCTT	654
OY		837	CTTCTCGAAGGGTTGAGGTTGAGAGACACTTCCATTGGAAGATACATCATTTGGA	896
Db		655	CTAGTTGAAGGACTAGATATAGGAGAGAAAAACAATGGAAGATCTCTTAGTAAC	714
OY		897	GCCACCAGGACGCCGATGAT---CTTGTTACTGACGGTGATTAAGGAACAATCACACTT	953
Db		715	GAGGATCATGAAGGTGATGATACCTTGAACCGTTGATGAGATGGAACAATGCGGCTT	774
OY		954	TATGGGCTGAAGAAAGGCCAGTCATGGAATGCTCGACCTTCTAAGGACCATCATGCTT	1013
Db		775	TATGGAACCCACGAAATCAATCGTACCTTGCTAGACCTGTCCAGAACAA-----	826
OY		1014	GGAAGTGTGCTTCTCTTGCACTCTCGTCATGGAGCATGTGAACAGAGTGTA	1073
Db		827	-ATAGCTCACTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATCTT	885
OY		1074	ATGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAAATATGCTCAAGCTGAGGA	1133
Db		886	AAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAATGCCAGAAGCAGGCGGA	945
OY		1134	AGTATGAGGAGCACATTTGTTCCAACTTTGGAAGTATGTTGAG-----TGTCA	1187
Db		946	AACACTCGGAGTGGGATTTTCCCTCATTTGGAAGCATGTTCACTACTGCGGATGCG	1005
OY		1188	CAGCATGCCAAAAATGAGCAGTGGGATGAAGAGA-----ATCTCATAGGATGAC	1238
Db		1006	CCTCACGGTAAACCGGCTCATTTGGAAAAAGACATPAGAGGCCATTACAACAAAGACAAT	1065
OY		1239	GAGGAGTACGCATCTGATGGTGCGAGAGGTGACTATGAG-----ACAATCTCAAT	1289
Db		1066	GATGACTATGCGACTGATGATGTGCGGGTGATGATGATGACTCGGACAACGATTTGCGT	1125
OY		1290	AGCCCATTTGCTGCCAGCAGGCAACAGGTGCGGAAGGAAGACATTTGTGACCATGCT	1349
Db		1126	AGCCCTTAATATGTGCGCGCCAGACCAACAGCATGGA---CAAGATATGATCCACATCTT	1182
OY		1350	CACCGTGGAAAGTCTTTGAGCATGAGAAAGCCAAACCTTTAGGGGAAGGTGAGATGCT	1409
Db		1183	ACAAGTGGAAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCA--AGGCAACGCGC	1239
OY		1410	GTCAGCAGCACTGATATCGGTGGGGGATGGCAGCTTGCTTGAAATGCTCAGAGAAGAA	1469
Db		1240	GAAAGTAGCATGGGAATGTGTGTGGCTTGCCATATGGGAT-----	1279
OY		1470	GCTGAGAAATGTTAGAAAGGAAGGTGCTTCAAAAGAGTCTACTTGCAACCAAGAGGAGTT	1529
Db		1280	-----ATAGATACGAAACGATGAATACAAAGAGTATTACTTAAAGAGATGGA	1329
OY		1530	CCTGGCTCAAGAAAGGGGCTCAAATTGTTCACTTCCCGGTGTGCGATGTTTTTGAGGCT	1589
Db		1330	GCTGAATCTCGCGGTGCTGATCTCTATTCCCGGAGGTCCGGATGCT-----GGA	1383
OY		1590	AGTGAAGTTTGTACATGCTGCTGCTTTAGTAAGTCAATCAGACACTTTTCTCAAAGGCTT	1649
Db		1384	GGCAGCTACATTCACGCTTCTGCCCCTTGTAAAGCAGATCTGTCTTGGTCTTAATCAGTT	1443
OY		1650	GCTGAACCAACGATGTCAAGATGCTGCCATGTTCAACCCCATCTGAAGGTAGCTGCCAAAGGT	1709

Db	1444	CATGGA-----TCCGCCATGGTTCCCCCGAGAAATTTGCTGCCTCTGGA	1488
QY	1710	TCACGTTGGAAGATTTGTTGGAACCTGGAGTGAGCGGTGCCCTGTATGTCGCTGTGGA	1769
Db	1489	CCACTCTGGTCTGCTCTTCTTGAAACCTGGGTGTTAAGCGTGCCCTGTGTTGTGTGTCGGC	1548
QY	1770	ATTCAGATCCCTTCAACAGTTTGTCTGGAATTAACGGTGTCTGTACTATACCCCACAATT	1829
Db	1549	ATTCAATATCTGCAGACAGTTTTCAGGTAATCAATGAGATTCTCTACTACATCTCAGATT	1608
QY	1830	CTTGAGCAAGCTGGTGTGGCAGTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCAGCA	1889
Db	1609	CTCGAACGGGCTGGCGTAGATATTCTTTCGAGCCCTCGGACTAAGTTCCATCTCTGGC	1668
QY	1890	TCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCTCCTGCAATTGGCTTTGCCATG	1949
Db	1669	TCATTCTCATCAGCGGTTTAACAAATTAATCAATGCTCCAGCCATTTGCTTGCCATG	1728
QY	1950	CTGCTTATGATCTTTCGGAAGAGTTTTCCTGCTAGGCACAAATTCAACTTGATA	2009
Db	1729	AGACTCATGATGTATCCGGAAGAGTCAATTACTTCTCTGGACAATCCAGTTCTCATTT	1788
QY	2010	GCATCTCTAGTATCCCTGGTTGTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCT	2069
Db	1789	GTTCTCACTTGTCTCTTGTCTCATCAGCGAGCTCATCCACATCAGCAAGTCGTGAACGA	1848
QY	2070	TTGCTCTCCACCCGTCAGTGTATTCTGCTTACTTCTGCTGCTTCTGTTATGGGATTGGTCCC	2129
Db	1849	GCACTCTCCACAGGTTGTGTGCTGCTTACTTCTGCTTCTTCTGTGATGGGTTACGGTCCC	1908
QY	2130	ATCCCAACATTTTATGTGACAGATCTTTCCAAACCAAGGTTCTGCGCTCTGTATTGCC	2189
Db	1909	ATTCMAACATCTCTGTTCTGAATCTTCCCAACAAGATCCGTTGTTCTTGATCGCC	1968
QY	2190	ATTTGTGCCCTTACATCTGTGATCGGAGATATCATCTGTCACTACAGCCCTTCTGTGATG	2249
Db	1969	ATATGTCTATGGTCTTTTGATGTGAGACATTAATGTGCACGTACTCACTTCCGTTCTC	2028
QY	2250	CTGAATGCTATTTGACTGGCGGGTGTTCACAGATATATGACAGTCTGATGCTTGAATTCC	2309
Db	2029	CTCAGCTCGATCGGACTAGTTGTGTTTACAGATTTACGCTGCGGTTGCGTTATCTCA	2088
QY	2310	TTTGTGTTGCTTCTCTTAAGTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACC	2369
Db	2089	TGGATCTTCGTTTACATGAAGTCCCGGAGACTAAAGGCATGCTTTGGAAGTTATCACA	2148
QY	2370	GAATTCTTTGCAATTGGTGC	2389
Db	2149	GACTACTTTGGCTTTGGAGC	2168

RESULT 7	
ABK51968	
ID	ABK51968 standard; cDNA; 1487 BP.
XX	
AC	ABK51968;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Wheat contig encoding Arabidopsis thaliana-like sugar transport protein.
XX	
KW	Wheat; Arabidopsis thaliana-like sugar transport protein;
KW	carbohydrate transport; grain filling; annual field crop;
KW	plant; gene; ss.
XX	
OS	Triticum aestivum.
XX	
FX	
FT	Key
FT	CDS
FT	
FT	
FT	

FT thaliana-like sugar transport protein"
/note="This sequence lacks a start codon"

FT US6383776-B1.
XX 07-MAY-2002.
PD 14-APR-1999; 99US-0291922.
XX 24-APR-1998; 98US-083044P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX WPI; 2002-451386/48.
DR P-PSDB; AAU97207.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 3; Column 47-50; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC representing a contig assembled from various wheat cDNA clones encodes
CC a portion of an Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other;

Query Match 24.9%; Score 691.2; DB 24; Length 1487;
Best Local Similarity 80.1%; Pred. No. 1.3e-153;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGGTGAGATGCTGTGAGCAGCAGTATCGGTGGGGATGGCAGCTTGCTGGAAA 1454
DB 12 GAGGGTGGGAGGCGAGTCAGCAGCAGCTGTGTTGGTGGGGGTGCACTCGCATGGAAA 71
QY 1455 TGGTCAGAGAGGAAGGTGAGATGTTGAGAAAGGAAGGTGTTCAAAAGAGTCTACTTG 1514
DB 72 TGGTCGAGCGCAGCAAGGCGAGGATGGCAAGAGGAGGCTTCAAAAGATCTACTTG 131
QY 1515 CACCAAGAGGAGTCTCTGCTCAAGAGGGGCTCAATTGTTCACTCCCGGTGGGC 1574
DB 132 CACCAAGAGGAGGAGTCCGAGCTCAAGAGGGGCTCTGTTGTTCACTCTCGTGGGGT 191
QY 1575 GATG---TTTGTAGGGTAGTGAAGTTGTACATGCTGCTGCTTTAGTACAGTACAGCA 1631
DB 192 GATGCCAGCAGAGGGGCGAGTGGTTTATACATGCTGCTGCTTGTGAAGCCACTCGGCT 251
QY 1632 CTTTCTCAAGGGTCTGTGTAACCAAGCATGTAGATGCT---GCCATGCTCACCCA 1688
DB 252 CTTTACTCCAAGGATCTTATGGAAGCGTATGGCCGCGGTCCAGCCATGATTCATCCA 311
QY 1689 TCTGAGTACGCTGCCAAGGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAGCGCT 1748
DB 312 TTGGAGGCGAGCTCCCAAGGTTCAATCTGGAAGATCTGTTGAACCTGCTGAGCGCT 371
QY 1749 GCCCTGTAGTCGGTGTGGAATTGAGATCTTCAACAGTTTGTGGAATTAACGGGTGT 1808
DB 372 GCATTGTTGTCGGTGTGGAATTGAGATCTTCAAGCTTTGCTGGAATTAATGAGATT 431
QY 1809 CTGTACTATACCCCAAAATTTCTGAGCAGCTGGTGGCAGTTAATCTTTCCAAATTT 1868
DB 432 CTCTACTATCTCTCAAAATTTCTGAGCAGCTGGTGGTGGTCTTCTTCCAAATCTT 491

QY 1869 GGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACTACTAATGCTT 1928
DB 492 GGCTTCAGTTCAGCATCAGCATCCATCTTGATCAGTTCTCTCACTACTACTAATGCTT 551
QY 1929 CCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1988
DB 552 CCAAGCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 611
QY 1989 GGCACAAATTCACATCTTGATAGCATCTGATTAATCTGCTGCTGCTGCTGCTGCTG 2048
DB 612 GGCACAAATTCACATCTTGATAGCATCTGATTAATCTGCTGCTGCTGCTGCTGCTG 671
QY 2049 TTGGGTACACTAGCCCATGCTTTGCTCTCCACCGTCAAGTATATCGTCTACTCTGCTGC 2108
DB 672 TTGAGTACGGGTGCCCCACGCTGTGCTCTCCACAGTTAGGTCATGTCTACTCTGCTGC 731
QY 2109 TTGCTTATGGGATTTGTGCTCCCATCCCAACATTTTATGTCAGAGATCTTCCAACCGAG 2168
DB 732 TTTGTCATGGGCTTTGGCCCATCCCAACATTTTATGTCAGAGATTTTCCCAACCGAG 791
QY 2169 GTTCGTGGCTCTGTATTTGCCATTTGTGCTCTTACATTTGAGATCGAGATATCATGCTC 2228
DB 792 GTCCGTGGTGTCTGCATGCTATTTGCGCCCTCACATTTGAGATTTGAGATATTTGTT 851
QY 2229 ACCTACAGCCTTCTGTGATGCTGAATGCTATTTGAGTGGCGGGTGTTCAGCATATAT 2288
DB 852 ACCTACAGCCTTCTGTGATGCTGAATGCTATTTGAGTGGCGGGTGTTCAGCATATAT 911
QY 2289 GCACTGTATGCTGTATTTCTTGTGTTGCTGCTTCTTAAAGTCCCTGAGACAAAGGG 2348
DB 912 GCAGTCGTTTGTGCTGATGCTTGTGTTGCTGCTTACCTAAAGTCCCAAGACAAAGGG 971
QY 2349 ATGCCCTTGAGGTTATTCAGAAATTTCTTGTGAGTGGTGGCAAGCAAGC--GGCTGCA 2405
DB 972 ATGCCCTTGAGGTTATTCAGAAATTTCTTGTGAGTGGTGGCAAGCAAGCAGGCCACC 1031
QY 2406 AAAGCTAATTTCTTGTGATCTTGTGCAACTATTCGACTG 2449
DB 1032 ATTGCTGATTCATCATGAGCTTTGTTTCAAGTTTGCACTG 1075

RESULT 8
ABX93204
ID ABX93204 standard; cDNA; 1487 BP.

AC ABX93204;
XX
DT 29-MAY-2003 (first entry)

DE cDNA encoding wheat sugar transport protein #2.
XX
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.

OS Triticum aestivum.

PN US2002178468-A1.

PD 28-NOV-2002.

PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.

PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;

XX O
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08332.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein
XX
PS Claim 2; Page 25-26; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other;

Query Match 24.9%; Score 691.2; DB 25; Length 1487;
Best Local Similarity 80.1%; Pred. No. 1.3e-153;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGTGGAGATGCTGTGAGCAGCACTGATATCGGTGGGGATGGCAGCTTGCTTGAA 1454
DB 12 GAGGTGGGGAGGAGCAGTACAGCACTGTATTTGGTGGGGGTGGCACTCGCATGAA 71
QY 1455 TGGTCAGAGAAGGAAGGTGAGATGGTGAAGAAAGGAAGGTGTTCAAAAGAGTCTACTG 1514
DB 72 TGGTCGAGCGCAGCAAGGCGAGATGGCAAGAAAGGAGGCTTCAAAAGATCTACTTG 131
QY 1515 CACCAAGAGGGAGTTCCTGGCTCAAGAGGGGCTCAATTGTTCACTTCCGGTGGTGC 1574
DB 132 CACCAAGAGGGGCTGGCCGACTCAAGAGGGGCTGTTGTTCACTTCCGGTGGTGC 191
QY 1575 GATG---TTTGGAGGTAGTGTGATGCTGCTGCTGCTTGAAGTCAAGTCAAGCA 1631
DB 192 GATGCCACGCAAGGGGAGTGGTGTATATCATGCTGCTGCTTGTGAAGCAGCTCGCT 251
QY 1632 CTTTCTCAAGGGTCTTGTGAACCAAGCATGTCAATGCT---GCCATGTTCAACCA 1688
DB 252 CTTTCTCAAGGATCTTATGGAAGCGTATGCGCGCCGTCAGCCATGATTCACA 311
QY 1689 TCTGAGTAGCTGCCAAGGTTCAAGTTGGAAAGATTTGTTGAACCTGAGTGAAGCGT 1748
DB 312 TTGGAGGCAAGCTCCCAAGGTTCAATCTGAAAGATCTGTTGAACCTGAGTGAAGCGT 371
QY 1749 GCCCTGTAGTGGTGTGGAATCAGATCCTTCACAGTTGCTGGAATTAACGGTGT 1808
DB 372 GCATGTGTGTCGGTGTGGAATCAGATGCTTCAGCAGTTGCTGGAATTAATGAGTT 431
QY 1809 CTGTACTATACCCCAAAATCTTGAAGCAAGCTGTGGGCAAGTTATTTCCAAATTT 1868
DB 432 CTGTACTATACCTCTCAAAATCTTGAAGCAAGCTGTGGGCTGTTCTTCTTCCAAATTT 491
QY 1869 GGTCTCAGCTCGCATCAGCATCCTTGTATCAGTCTCTCACTACCTACTAATGCTT 1928
DB 492 GGGCTCAGTTCAGCATCAGCATCCTTGTATCAGTCTCTCACTACCTACTAATGCTT 551
QY 1929 CCTTCATTTGGCTTGGCATGCTGCTTATGATCTTCCGGAAGAGGTTTTCGCTGA 1988
DB 552 CCAAGCATTTGGTGTAGCCATGAGACTTATGATATATCTGGAAGAGGTTTCTGCTACTG 611
QY 1989 GGCACAATTTCAATCTGTATAGCATCTAGTATCTCTGCTGTGTCCAATCAATGAT 2048
DB 612 GGCACAATTTCCATCTGTATAGCATCTTATGTTTGGGTGTGTCATGTTATCAAC 671

QY 2049 TTGGTACACTAGCCCAAGCTTTGCTCTCCACCGTCAGTGTATCGTCTACTCTGCTGC 2108
DB 672 TTGAGTACAGGTGCCCAAGCTGTGCTCTCCACAGTTAGCGTCATGTCTACTCTGCTGC 731
QY 2109 TTGCTATGGATTTGGTCCCATCCCAACATTTATGTGCAGAGATCTTCCCAACGAG 2168
DB 732 TTTGTATGGGCTTTGGCCCCGATCCCAACATTTATGTGCAGAGATTTTCCCAACGAG 791
QY 2169 GTTCGTGGCTCTGTATGGCATTTGTGCTTTCATTTCTGATCGAGATATATCGTC 2228
DB 792 GTCCGTGGTGTGTGATCGCTATTTGGCCCTCAATCTGATTTGTGACATTAATGTT 851
QY 2229 ACCTACAGCCTTCTGTGATGCTGAATGCTATTTGAGCTGCGGGTGTTCAGCATATAT 2288
DB 852 ACCTACAGCCTGCTGTGATGCTGAATGCTATTTGATGCTGCGGGTGTTCGATATAT 911
QY 2289 GCAGTCGATGCTGAATTTCTTGTGCTGCTTCTTAAGTCCCTGAGACAAAGGG 2348
DB 912 GCAGTCGCTTGTGCTGATGCTTGTGCTTGTGCTTACCTAAAGTCCAGACAAAGGG 971
QY 2349 ATGCCCCCTGAGGTATTAACCGAATCTTTGCAAGTTGCTGCGAAGCAAGC---GGCTGCA 2405
DB 972 ATGCCCCCTGAGGTATTAACCGAATCTTTGCAAGTTGCTGCGGTTGGGCGAAGCAAGCGCAGCCACC 1031
QY 2406 AAGCCTAATTTCTTGTGATCCTTGTGCTGCAACTATTCACCTG. 2449
DB 1032 ATTGCTGATTCATCATGAGGCTTGTGTTTTCAGTTTGCAACACTG 1075

RESULT 9
ABK51966
ID ABK51966 standard; cDNA; 1692 BP.

AC ABK51966;
DT 27-AUG-2002 (first entry)
XX

DE Soybean cDNA clone ssl.pk0022.fl encoding sugar transport protein.

KW Soybean; Arabidopsis thaliana-like sugar transport protein;

KW carbohydrate transport; grain filling; annual field crop;

KW plant; clone ssl.pk0022.fl; gene; ss.

OS Glycine max.

XX
FH Key location/Qualifiers

FT CDS 9..1469

FT /tag= a

FT /partial

FT /product= "portion of a soybean Arabidopsis

FT thaliana-like sugar transport protein"

FT /note= "This sequence lacks a start codon"

XX
PN US6383776-B1.

XX
PD 07-MAY-2002.

XX
PF 14-APR-1999; 99US-0291922.

XX
PR 24-APR-1998; 98US-083044P.

XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;

XX
DR WPI; 2002-451386/48.

XX
DR P-PSDB; AAU97205.

XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for

PT preparing transgenic plants with altered carbohydrate distribution -

XX
PS Claim 3; Column 41-44; 54pp; English.

CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence encodes
CC a portion of a soybean Arabidopsis thaliana-like sugar transport
CC protein.

SD Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 other;

Query Match 21.4%; Score 594.6; DB 24; Length 1692;

Best Local Similarity 66.18; Pred. No. 1e-130;

Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;

QY	929	CGGTGATGAAGAACAAATCACTTATATGGGCGCTGAAGAGGCCAGTCATGATTGCTCG	988
Db	17	CAGAGAAAAAGATCAAAATTAAAGTTGTATGACCAGAAACAAGCCAGTCCGTGGTGTAG	76
QY	989	ACCTTCTAAGGAGCCCAATCATGCTTGAAGTGTCTTCTTGCATCTCGTCAAGGAG	1048
Db	77	ACCTGTTGCTGGACCAAAATCTGTTGGC-----CTTGATCTAGAAAGGAAG	124
QY	1049	CATGTTGAACCAAGTGAACCCCTTATGATCCGATTGTGACACTTTTGGTAGTGTCCA	1108
Db	125	CATGGCAAAATCCAAAGCAGT---CTAGTGAGCCCTCTAGTGAACCCCTTTGGTAGTGTACA	181
QY	1109	TGAGAAATATGCCCTCAAGCTGGAAGGAAGTATGAGGAGCACTTGTTTCCAAACTTTGGAAG	1168
Db	182	TGAGAACTCCCAAGAACAGGA-----AGCACCCCTTTTCCACACTTTGGGAG	229
QY	1169	TATGTTCAGTGTCAAGATCAGCATGCCCCAAAATGAGCAGTGGATGAAGAGAAATCTTCA	1228
Db	230	TATGTTCAGTGTGGGGAAATCAGCCAAAGAAATGAAGATTGGATGAGAAAGCCTAGC	289
QY	1229	TAGGATGACGAGAGTACGCATCTGATGCTGCAGAGGTGACTATGAGACAATCTCCA	1288
Db	290	CAGAGAGGGTGATGATTATGTCTCTGATGCT-----GGTGAATCTGATGACAAATTTGCA	343
QY	1289	TAGCCCATTTGCTGTCCAGGAGGCCAACAGGTGCGGAAGGGAAGACATTTGTCACCATGG	1348
Db	344	GAGTCCATTGATCTCAGTCAAAACAAGAGTCTGGA---TAAGACATACCTCTCATGTC	400
QY	1349	TCAACGTTGAAGTGTGATGACATGAGAAGGCAAAACCTCTTAGGGGAGGCTGAGATGG	1408
Db	401	CCATAGTAACCTTGCAAGCATGAGGCAAGGTAGTCTTTACATGGAATTCAGAGAGAAC	460
QY	1409	TGTGAGCAGCACTGATATCCGTGGGGGATGGCAGCTTGTGGAATGTCAGAGAGAGA	1468
Db	461	CACTGGTAGTACTGGGATTGCTGTGCTGGCAAGCTAGCATGGAATGTCTGAAGAAGA	520
QY	1469	AGGTGAAGATGTAGAAAGGAAGGTGCTTCAAAAGAGTCTACTTGCAACCAAGAGGAGT	1528
Db	521	GGGCCAGATGGAAGAGGAAGGTGCTTCAAGAGATATATTACACCAAGATGTGG	580
QY	1529	TCCTGGCTCAAGAAAGGGCTCAATTGTTTCACTTCCCGGTGGTGCGATGTTTTGAGGG	1588
Db	581	TTCTGATCTAGACGTGGGTCTGTGTTTCACTCCC---TGGCGGTGATTTACCACTGA	637
QY	1589	TAGTGAGTTTGTACATGCTGTGCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGCT	1648
Db	638	CAGTGAGTTGTACAGGCTGTCTCTGTGTGAGTCAAGCTTCCCTTTATATAGAGACT	697
QY	1649	TGCTGAACCAAGCATGTCAAGTCTGCTGCCATGTTTACCACTGAGGTAGCTGCCAAAG	1708
Db	698	TATGCGTCAACGGCCAGTTGACCAAGTATGATTATCCCTCTGAAACAATTGCAAAAG	757
QY	1709	TTCAAGTTGAAAGATTGTTTGAACCTGAGTGAAGCGTGCCTGTTAGTGTGTTGG	1768

Db	758	GCCAAGTTGAGTGAATCTTTTGAACCTGGGGTGAAGCATGCATTGATGTGGGGGTGGG	817
QY	1769	AATTCAGATCCTTCAACAGTTTGTGTAATAACGGTGTCTGTACTATACCCCAAAAT	1828
Db	818	AATGCAATTTCTTCAGCAGTCTCTGGTATAAATGGGGTCTCTACTATACGCTCAAAAT	877
QY	1829	TCTTGAGCAAGCTGGTGTGGCAGTTATTTCTTTCCAAATTGGTCTAGCTCGGCATCAGC	1888
Db	878	TCTTGAGCAGGACGAGGTGTGGTTATCTTCTTTCAGCCTAGGCGTTGGTTCCTACTTCTTC	937
QY	1889	ATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCTGCAATTGGCTTTGCCAT	1948
Db	938	ATCCTTCTTATATAGTGGCGGTGACAACTGTGTGATGCTTCTGTATAGCCATGGCCAT	997
QY	1949	GCTGCTTATGATCTTTCGGAGAAGGTTTTCGTGCTAGGACACAATTCCAATCTTGAT	2008
Db	998	GAGGCTCATGATATTTTCAGCAGACAGAAGCACTTGTGCTCAGTACAAATCCCGTCTAAT	1057
QY	2009	AGCATCTCTAGTATCTCTGCTGTGTGTCCAATCTAATTGATTTGGGTACACTAGCCCATGC	2068
Db	1058	AGCAGCTCTTCTCATATTAAGTCTCTGGAGAGTCTGTGGAATTTGGGATCCACTGCAAAATGC	1117
QY	2069	TTTGTCTCTCCACCGTCAGTGTATTGCTCTACTTTCGCTGCTTGTATNGGATTTGGTCC	2128
Db	1118	ATCAATCTCAACCAATTAAGTGTATTGTCTATTTCTGTTCCTTGTCAATGGATTTGGACC	1177
QY	2129	CATCCCCCAACATTTTATGTGCAGAGATCTTTCCAACCAAGGTTGCTGAGCTCTGTATTC	2188
Db	1178	AATTCCTAATATACTTGTGACAGAGATCTTCCCACTCGAAGTTCGTGCTCTGCATTC	1237
QY	2189	CATTGTGCCCTTACATTTCTGGATCGGAGATATCATCGTCACCTACAGCCTTCTGTGAT	2248
Db	1238	TATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTCACTACACACTCCACAGTTAT	1297
QY	2249	GCTGAATGCTATTTGAGCTGGCGGGTGTTCAGCATATATGCAGTGTATNGCTTGAATTC	2308
Db	1298	GCTCAATTCGTAGGCTCGCTGTGTTTGTGTAATTTATGCTGTGCTGTCTTCATAGC	1357
QY	2309	CTTTGTGTGCTTCTTCTTAAAGTCCCTGAGACAAAGGGATGCCCTTGAGGTTATTAAC	2368
Db	1358	ATGGGTGTTGTCTTTTGAAGTTCAGAAACCAAGGCATGCCACTGGAAGTATCAT	1417
QY	2369	CGAATTCCTTGCAGTTGGTGGCAAGCA	2395
Db	1418	TGAGTTCCTTCTGTGCGAGCAAAAACA	1444

RESULT 10

ABX93202

ID ABX93202 standard; cDNA; 1692 BP.

AC ABX93202;

DT 29-MAY-2003 (first entry)

DE CDNA encoding soybean sugar transport protein #2.

Arabidopsis thaliana-like sugar transport protein: corn: rice: wheat:

carbohydrate distribution: plant: gene: ss

Glycine max.

XX
XX
DN
IIS2002178468-A1

XX 20 NOV 2003

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

XX

PR 14-APR-1999; 99US-0291922.

PÀ (ALLE/) ALLEN S M.

PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08330.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein
XX
XX Claim 2; Page 22-23; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 other;

Query Match 21.4%; Score 594.6; DB 25; Length 1692;
Best Local Similarity 66.1%; Pred. No. 1e-130;
Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;

QY 929 CGGTGATAAGGAACAATCACACTTATGGGCGCTGAGAAGGCCAGTCATGATTGCTCG 988
Db 17 CAGAGAAAAAGATCAATTAAGTTGTATGACCAGACAGGCCAGCTCTGGTGTCTAG 76
QY 989 ACCCTCTAAGGGAGCCCATCATGCTTGAAGTGTGCTTCTCTTGATCTCTCATGGAG 1048
Db 77 ACCGTGTGTGACCAAAATTCGTGTGC-----CTTGATCTAGGAAAGGAG 124
QY 1049 CATGTGAACCAAGAGTGTACCCCTTATGATCCGATGTGACACTTTTGTAGTGTCCA 1108
Db 125 CATGGCAATCCAAAGCAGT--CTAGTGACCCTCTAGTGACCTCTTTGGTAGTGTACA 181
QY 1109 TGAGAAATATGCTCTCAAGCTGAGAGAAATATGAGACACATTTGTTCCAACTTTGGAAG 1168
Db 182 TGAGAGCTCCCAAGAACAGCA-----AGCACCCCTTTTCCACACTTTGGGAG 229
QY 1169 TATGTTAGTGTCAACAGATCAGCATGCCAAAATGAGCAGTGGGATGAAGAATCTTCA 1228
Db 230 TATGTTAGTGTGGGGAATCAGCCCAAGAAATGAGATGGGATGAAGAAAGCCTAGC 289
QY 1229 TAGGATGACGAGAGTAGCATCTGATGTGTGACAGAGTGACTATGAGACAATCTCCA 1288
Db 290 CAGAGAGGTGATGATTATGTCTGTGATCT-----GGTGATCTGATGACAATTGTCA 343
QY 1289 TAGCCCATTTGCTGTCCAGGCAAGGAGTGGGAGGGAAGGACATTGTGACCATGG 1348
Db 344 GAGTCCATTTGATCTCAGCTCAAAACAGAGTCTGGA--TAAGGACATACCTCCTCATGC 400
QY 1349 TCACCGTGAAGTGTCTTGAAGCATGAGAAGCAAAACCTCTTAGGGGAGGGGTGAGATGG 1408
Db 401 CCAATGTAACCTTGCAAGCATGAGGCAAGTAGTCTTTTACATGGAATTCAGAGAAACC 460
QY 1409 TGTGACGACACTGATATCGTGGGGGATGGCAGCTTGCTTGGAATGCTCAGAGAAGA 1468
Db 461 CACTGTGTACTGGGATTTGGTGTGGCAGCTAGCATGGAATGCTTGAAGAAGA 520
QY 1469 AGGTGAAATGTTAGAAAGAGAGTGTGTTCAAAAGAGTCTACTTGCACCAAGAGGAGT 1528
Db 521 GGGCCCATGATGAAAGAGAGAGGTGGCTTCAAGAGATATATTTACCAAGATGTTGG 580

QY 1529 TCCTGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGGTGCGATGTTTGAAGG 1588
Db 581 TTCTGATCTAGACGTGGGTCTGTGTTTCACTCCC--TGCGGTGATTTACCACTGA 637
QY 1589 TAGGAGTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTTTTCAAGGCTCT 1648
Db 638 CAGTGAGGTGTACAGGCTGCTGCTCTGTGAGTCAGCTGCCCTTATATAGGACCT 697
QY 1649 TGCTGAACCAAGCATGTCAGATGCTGCCATGTTCAACCATCTGAGGTAGTGCCTGAAG 1708
Db 698 TATGCTCAACGGCCAGTTGAGACCAAGCTATGATTCAATCCCTGAAACAAATTGCAAAAG 757
QY 1709 TTCACTTGGAAGATTGTTTGAACCTGAGTGAAGGCGTCCCTGTAGTCGGTGTGG 1768
Db 758 GCCAAGTTGAGTGAATCTTTTGAACCTGGGGTGAAGCATGCAATTGATGTGGGGTGG 817
QY 1769 AATTGATCTTCAACAGTTTGTGTAATAAACGGTGTCTGTACTATACCCCAAAAT 1828
Db 818 AATGCAATTTCTCAGCAGTTCTGTGTATAAATGGGCTCTACTATACGCTCAAAAT 877
QY 1829 TCTGAGCAAGCTGTGTGAGCAGTTATTTCTTCCAAATTGCTCTCAGCTCGGCATCAGC 1888
Db 878 TCTTGAAGCAGCAGGTGTGTGTATCTTCTTCAAGCCTAGGCTGTGTACTTCTTTC 937
QY 1889 ATCCATCTGATCAGTTCTCTCACTACCTTAATAGCTTCTTGCATTTGGCTTGGCAT 1948
Db 938 ATCTTTCTTATTAGTGGGTGACCAACCTGTGTGATGCTTCTGTATAGCCATGCCAT 997
QY 1949 GCTGCTTATGATCTTTCGGAAGAAGTTTTGTCTGCTAGGCACAAATTCATTTGAT 2008
Db 998 GAGGCTCATGATATTTCAGGACAGAAGACTTTGCTGCTCAGTACAAATCCCGTCTAAT 1057
QY 2009 AGCATCTTATGATTAATCTGTGTGTGTCTCAATTAATGATTTGGGTACACTAGCCATGC 2068
Db 1058 AGCAGCTCTTCTCATATTAGTCTCTGGAAGCTTTGTGATTTGGATCCACTGCAAAATGC 1117
QY 2069 TTTGCTCTCCACCGTCAAGTGTATGCTACTTCTGCTGCTTGTATAGGATTTGGTCC 2128
Db 1118 ATCAATCTCAACCATTAAGTGTATGTCTATTTCTGTTCTTGTATGAGATTTGACC 1177
QY 2129 CATCCCAACATTTTATGTCAGAGATCTTTCCAAACAGGGTGTGAGCTGTATATGC 2188
Db 1178 AATTCCTAATATACCTTGTGCAAGATCTTCCCACTCGAGTCTGTGCTGTGCAATGC 1237
QY 2189 CATTTGCTTATCATCTGTGATCGGATATCATCTCACCTACAGCCTTCTGTGAT 2248
Db 1238 TATTTGTCCCTTACCTTTTGGATCTGTGATATCATTTGTCACTACACACTCCAGTTAT 1297
QY 2249 GCTGAATGCTATGTGACTGGGCGGTGTTTACAGCATATATGCAAGTGTATGCTTATTC 2308
Db 1298 GCTCAATTTGTAGGCTGCTGTGTTTGTGTTTATGCTGTGCTGTCTTCAATAGC 1357
QY 2309 CTTGTGTTGCTTCTTAAAGTCCCTGAGACAAAGGGATGCCCTTGAGTTATTTAC 2368
Db 1358 ATGGGTGTTGCTTTTGAAGTTCAGAAACCAAGGGCATGCGACTGGAAGTATCAT 1417
QY 2369 CGAATCTTTGAGTGTGCGGAAGCA 2395
Db 1418 TGAGTTCTTCTGTGAGCAAAACA 1444

RESULT 11
ABK51969
ID ABK51969 standard; cDNA; 1009 BP.
XX
AC ABK51969;
XX
DT 27-AUG-2002 (first entry)
XX
DE wheat cDNA clone wreln.pk0006.b4 encoding sugar transport protein.
XX wheat; Arabidopsis thaliana-like sugar transport protein;
KW

KM carbohydrate transport; grain filling; annual field crop;
KM plant; clone wreln.pk0006.b4; gene; ss.
OS Triticum aestivum.
XX
XX
FH Key Location/Qualifiers
FT CDS 2..688
FT /*tag= a
FT /partial
FT /product= "Portion of a wheat Arabidopsis
thaliana-like sugar transport protein"
FT /note= "this sequence lacks a start codon"
XX
XX PN US6383776-B1.
XX
XX PD 07-MAY-2002.
XX
XX PF 14-APR-1999; 99US-0291922.
XX
XX PR 24-APR-1998; 98US-083044P.
XX
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
XX DR WPI; 2002-451386/48.
XX P-PSDB; AAU97208.
XX
XX PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution -
XX
XX PS Claim 3; Column 51-54; 54pp; English.
XX
XX CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence encodes
CC a portion of a wheat Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 1009 BP; 237 A; 233 C; 234 G; 305 T; 0 other;

Query Match 14.2%; Score 394.4; DB 24; Length 1009;
Best Local Similarity 74.4%; Pred. No. 2.3e-83;
Matches 497; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 1730 TGAACCTGAGTGAGGCGCTGCCCTGTAGTCGGTGTGGAATTCAATCCTTCAACAGTT 1789
Db 1 TGAACCTGAGTGAGCATGCACTGTTCGTTGGCATAGGATTAACAGATCCTGCAGCAGTT 60
QY 1790 TGCCTGAATTAACGGGTGTTCTGTACTATACCCCAAAATTCTTGAGCAAGCTGTGCGC 1849
Db 61 TCGGGGTATCAATGAGTCCCTCTACTACACACCTCAGATACTTGAGCAAGCAGGTGTCGG 120
QY 1850 AGTTATCTTTCCAAAATTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCT 1909
Db 121 GGTTCCTCTATCAAAACATGGAATGAGCTCTTCTCCTCAGCATCTATCTTATTAGTGCCTT 180
QY 1910 CACTACCTTACTAATAGCTTCCTTCGATCTGGCTTGGCATGCTGCTTATGATCTTTCCGG 1969
Db 181 GACAACCTTGCTGATGCTCCCAAGCATGGCATCGCCATGAGACTCATGATATGTCAAG 240
QY 1970 AAGAAGTTTTTGTGCTAGGACAAATTCATCTTGATAGCATCTCTAGTATCCGTGT 2029
Db 241 AAGAAGTTTCTTCTCTCTTCAACATCCCTGTCTGTAGTAGCGCTAGCTGTCTGTGT 300
QY 2030 TGTGTCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTGCTCTCCACCGTCAGTGT 2089

Db 301 TTTAGTAATGTTCTGGATGTCGGAACCATGTTGACCGCTCGCTCTCAACGATCAGCGT 360
QY 2090 TATCGTCTACTTCTGCTGCTCTGTTATGGAATTTGCTCCCATCCCAACATTTATGTGC 2149
Db 361 CATGCTATTTCTGCTTCTTCTGTCATGCGGTTTGGGCTTATCCCAATATTTCTCTGCGC 420
QY 2150 AGAGATCTTTCCAAACGAGGTTGCGCTCTGTATTTGCCATTTGTGCTTTACATTTCTG 2209
Db 421 GGAGATTTTCCCACTCTGTCCGTGSCATCTGCATAGCCATCTGCGCGCTAACCTTCTG 480
QY 2210 GATCGAGATATCATGCTCAACCTACAGCCTTCCCTGTGATGCTGAATGCTATTTGACTGCGC 2269
Db 481 GATCGCGACATCATGCTGACATACACTCTCCCGGATGCTCAATGCCATTGCTCTCGC 540
QY 2270 GGGTGTTCAGCATATATGCACTGCTATGCTTGAATTTCTTGTGTTGCTTCTCTTAA 2329
Db 541 TGGAGTCTTCGGCATATATGCCATCGTTGTGTACTAGCCCTTGTATTTGCTTACATGAA 600
QY 2330 GGTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTAACGAATCTTTGAGTTGCTGC 2389
Db 601 GGTCCCTGAGACAAAGGGCATGCCCCCTGAGGTCATCACCGAGTCTTCTGTGCGGGGC 660
QY 2390 GAAGCAAG 2397
Db 661 AAAGCAGG 668

RESULT 12
ABX93205
ID ABX93205 standard; cDNA; 1009 BP.
XX
XX AC ABX93205;
XX
XX DT 29-MAY-2003 (first entry)
XX
XX DE cDNA encoding wheat sugar transport protein #3.
XX
KM Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KM plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Triticum aestivum.
XX
XX PN US2002178468-A1.
XX
XX PD 28-NOV-2002.
XX
XX PF 17-JAN-2002; 2002US-0051902.
XX
XX PR 24-APR-1998; 98US-083044P.
XX PR 14-APR-1999; 99US-0291922.
XX
XX PA (ALLE/) ALLEN S M.
XX PA (HITZ/) HITZ W D.
XX PA (KINN/) KINNEY A J.
XX PA (TING/) TINGEY S V.
XX
XX PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
XX DR WPI; 2003-340957/32.
XX DR P-PSDB; ABU08333.
XX
XX PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein -
XX
XX PS Claim 2; Page 27; 56pp; English.
XX
XX CC The present invention relates to the isolation of Arabidopsis
XX thaliana-like or Beta vulgaris-like sugar transport proteins, and
XX the polynucleotide sequences encoding them. The plant sugar transport
XX proteins of the invention have been isolated from corn, rice, soybean,
XX and wheat. The polypeptides of the invention may be used for altering

CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.

XX
SQ Sequence 1009 BP, 237 A; 233 C; 234 G; 305 T; 0 other;

Query Match 14.2%; Score 394.4; DE 25; Length 1009;
Best Local Similarity 74.4%; Pred. No. 2.3e-83;
Matches 497; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

```
QY 1730 TGAACCTGAGTGAAGCGGCTGCTGTTAGTGGTGTGAATTGAGATCCTTCAACAGTT 1789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 TGAACCTGAGTGAAGCATGACATGTTGCTGGCATAGATTACAGATCCTGCAGCAGTT 60

QY 1790 TGCCTGAATAAAGCGTGTCTGTACTATACCCACAATTCTTGAGCAAGCTGTGTCG 1849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TCGGGTATCAATGAGAGTCCCTCTACTACACACCTCAGATACTTGAGCAAGAGTGTGCG 120

QY 1850 AGTTATCTTCCCAATTTGCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTT 1909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GGTCTTCTATCAAAACATGGACATAAGCTCTTCTCAGCATCTATTCTTATAGTCCCT 180

QY 1910 CACTACCTTAATAATGCTCTCCTTGCAATGGCTTGGCATGCTGCTTATGATCTTCCG 1969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GAGAACCTTGCTGATGCTTCCAGCATGGCATCGCCATGAGACTCATGATATGTCAGG 240

QY 1970 AAGAAGTTTTTGTGCTGCTAGGACAAATCCAACTGTGATAGCATCTCTAGTTATCCTGT 2029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AAGAAGTTTTTCTTCTCTCTTCAACAATCCCTGTTGATAGAGGCTAGCTGTCTGTGT 300

QY 2030 TGTGTCAATCTAATTGATTGGGTACACTAGCCCATGCTTGTCTCCACCGTCAGTGT 2089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TTATGATGAATGTTCTGATGTGGAAACATGGTGCACGTGCGCTCTCAACGATCAGCGT 360

QY 2090 TATCTCTACTTCTGCTGCTTCTGTAATGGGATTGTGCCATCCCAACATTTATGTGC 2149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 CATGCTCTATTCTGCTTCTGCTCATGGGGTTTGGCCCTATCCAAATATTCTCTGCGC 420

QY 2150 AGAGATCTTCCCAACAGGGTCTGCGCTCTGTATGTCATTTGCTTACATTTCTG 2209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 GAGATTTTCCCAACCTCTGTCGGTGCATCTGCATAGCCATCTGCGCTAACCTTCTG 480

QY 2210 GATCGAGATATCATGCTCACCCTTACAGCTTCTGTGATGCTGAATGCTATTGACTGCG 2269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 GATCGCGACATCATGCTGACATACCTCTCCCGATGCTCAATGCTGATGCTCTGCG 540

QY 2270 GGGTGTCTTTCAGCATATATGACATGCTATGCTGATTTCTTGTGCTTCTCTTAA 2329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 TGGAGCTTTCGGCATATATGCAATGCTTGTGTAAGCTTGTATGCTTACATGAA 600

QY 2330 GGTCCCTGAGACAAAGGGATGCCCTTGAGGTTATTACCGAATCTTTGAGTTGTCG 2389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 GGTCCCTGAGACAAAGGGATGCCCTTGAGGTTATTACCGAATCTTTGAGTTGTCG 660

QY 2390 GAAGCAAG 2397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 AAAGCAGG 668
```

RESULT 13

ABL71325 standard; cDNA; 282 BP.

XX ABL71325;

XX 14-MAY-2002 (first entry)

DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:699.

KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.

XX Zea mays.

XX US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-0294093.

XX 21-APR-1998; 98US-082567P.

XX (LALG/) LALGUDI R V.

XX (ITOL/) ITO L Y.

XX (SHER/) SHERMAN B K.

XX Lalgudi RV, Ito LY, Sherman BK;

XX WPI; 2002-163647/21.

PT Novel purified corn tassel-derived polynucleotide useful for
PT determining altered gene expression, to recover regulatory elements and
PT breeding programs -

PS Claim 1; SEQ ID 699; 201pp; English.

XX The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful
CC in the evaluation, and alteration of desired characteristics associated
CC with growth and development, disease resistance, environmental
CC adaptability, quality and yield, and as molecular markers for studying
CC inheritance of multigene traits in a plant breeding program. (I) can be
CC used to produce a tassel-specific profile of gene transcription, a
CC transcript image, to clone regulatory elements for use in transformation
CC vectors, to express a polypeptide, to identify, isolate or extend
CC identical or related corn tassel nucleic acid sequences from DNA
CC libraries, in nucleic acid hybridisation or amplification technologies,
CC as query sequences to determine homology of known sequences, as probe
CC for use in Southern or Northern hybridisation, and to identify the
CC presence of and/or to determine the degree of similarity between two
CC (or more) nucleic acid sequences.

XX Sequence 282 BP; 54 A; 63 C; 61 G; 95 T; 9 other;

Query Match 9.3%; Score 257.2; DB 24; Length 282;
Best Local Similarity 95.4%; Pred. No. 4.7e-51;
Matches 270; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

```
QY 1938 GGCTTGGCCATGCTGTTATGATCTTCCGGAAGAGTTTGTGCTGCTAGGCAATT 1997
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GGCTTGGCCATGCTGTTATGATCTTCCGGAAGAGTTTGTGCTGCTAGGCAATT 60

QY 1998 CCAATCTGATAGCATCTAGTTATCCCTGTTGTGCCAATCTAATTGATTGGGTACA 2057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CCAATCTGATAGCATCTAGTTATCCCTGTTGTGCCAATCTAATTGATTGGGTACA 120

QY 2058 CTAGCCCATGCTTGTCTCCACCGCTCAGTGTATTGCTACTTCTGCTGCTGTTATG 2117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CTAGCCCATGCTTGTCTCCACAGTCAAGTGTATTGCTACTTCTGCTGCTGTTATG 180

QY 2118 GGAATTGGTCCCATCCCAACATTTTATGTGCAGACATCTTCCAACACAGGGTTCGTGC 2177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GGAATTGGTCCCATCCCAACATTTTATGTGCAGACATCTTCCAACACAGGGTTCGTGC 240
```

QY 2178 CTCTGATGTGCCATTGTGCCCTTACATCTCGATCGAGATA 2220
|||||
Db 241 CTCTGNATGTCCATNTG-GGCTTACATNNNGANNGANATA 282
RESULT 14
ABK51964
ID ABK51964 standard; cDNA; 870 BP.
XX ABK51964;
AC
XX 27-AUG-2002 (first entry)
DT
XX
DE Rice cDNA clone rl86.pk0003.d5 encoding sugar transport protein.
XX
KW Rice; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant; clone rl86.pk0003.d5; gene; ss.
XX
OS Oryza sativa.
XX
FH Key location/Qualifiers
FT CDS 8..403
FT /*tag= a
FT /partial
FT /product= "Portion of a rice Arabidopsis
thaliana-like sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
PN US6383776-BL.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-0291922.
XX
PR 24-APR-1998; 98US-083044P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-451386/48.
DR P-PSDB; AAU97203.
XX
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution
XX
XX
PS Claim 3; Column 33-36; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
CC transport proteins in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence encodes
CC a portion of a rice Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;
Query Match 8.2%; Score 227; DB 24; Length 870;
Best Local Similarity 74.9%; Pred. No. 1e-43;
Matches 284; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 2020 TTATCCGTGTGTGCCAATCTAATGATTGGGTACACTAGCCCATGCTTGTCTCCA 2079
|||||
Db 12 TAACCTGATCTGTGCATATCTCGATGTGGGACCATGTTCTACCTCACTGTCCA 71
CCGTACGTGTTATCGTCTACTTCTGTGCTTGTATGGGATTTGGTCCCATCCCAACA 2139
|||||

Db 72 CAGTCAGTGCATACTCTACTTCTGCTTCTTTGTCAATGGGGTTGGGCCATTCCAAACA 131
QY 2140 TTTATGTGCAGAGATCTTTCCAAACGAGGTTGTCGCTCTGTATTGCCATTGTGCTT 2199
|||||
Db 132 TTCTCTGTGCAGAGATTTCCCGACCAACCGTTGTCGTCATGTGCATAGCCATCTGTGCC 191
QY 2200 TTACATCTGTGATCGGAGATATCATCGTCAACCTACAGCCTTCTGTGATGCTGAATGCTA 2259
|||||
Db 192 TAACATCTGTGATCGGTGATATCATGTGACATACACCTCCCGTGATGCTCAACGCCA 251
QY 2260 TTGCACTGGCGGGGTGTTTCAACCATATATGACAGTGTGATGCTTGTGTTG 2319
|||||
Db 252 TTGCACTGCTGTGAGTGTGGAATCTACGCAAGTGTCTGCATACTGCTTCTGTGTTG 311
QY 2320 TCTTCCTTAAGTCCCTGACACAAAGGGATGCCCTTGAGGTTATTACCGAATCTTTG 2379
|||||
Db 312 TCTTCATGAAGGTGCCGAGACAAAGGCATGCTCTTGAAGTCATCACCGAGTTCTTCT 371
QY 2380 CAGTTGTGCGAAGCAAGC 2398
Db 372 CTGTGCGAGCAAAAGCAGGC 390
RESULT 15
ABX93200
ID ABX93200 standard; cDNA; 870 BP.
XX
AC ABX93200;
XX
DT 29-MAY-2003 (first entry)
XX
XX cDNA encoding rice sugar transport protein #2.
DE
XX
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Oryza sativa.
XX
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-0051902.
XX
PR 24-APR-1998; 98US-083044P.
PR 14-APR-1999; 99US-0291922.
XX
XX (ALLEN/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08328.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein
XX
PS Claim 2; Page 18; 56pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis
CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
CC the polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering
CC the level of expression of a sugar transport protein in a host cell,
CC by transforming a host cell with a chimeric construct encoding all,
CC or a portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to

CC control carbohydrate transport and distribution in plants.
CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.

XX
SQ Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;

Query Match 8.2%; Score 227; DB 25; Length 870;
Best Local Similarity 74.9%; Pred. No. 1e-43;
Matches 284; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
QY 2020 TTATCCGTGTTGTGTCATTAATTGATTTGGGTACACTAGCCCATGCTTGTCTCCA 2079
    |||||
Db 12 TAACCTTGATCTGTCAATATTCGATGTGGGACCATGTTCTAGCCTCACTGTCCA 71
    |||||
QY 2080 CCGTCACTGTTATCGTCTACTTCTGCTGCTTGTATGAGATTGTCATCCCAACA 2139
    |||||
Db 72 CAGTCAGTGCATACTACTTCTGCTTCTTGTATGAGGTTCCGCTATTCGAACA 131
    |||||
QY 2140 TTTTATGTCAGAGATCTTCCCAACGAGTTGCTGCTCTGATTTGCCATTGTGCT 2199
    |||||
Db 132 TTCTCTGTGACAGATTTTCCCGACCACTGCTGTCATCTGATAGCCATCTGTGCC 191
    |||||
QY 2200 TTACATTTGGATCGAGATATCATCTACCTACAGCCTTCTGTGATGCTGAATGCTA 2259
    |||||
Db 192 TAACATTTGATCGGTGATATCATGTGACATACACCTCCCGATGCTCAACGCCA 251
    |||||
QY 2260 TTGACTGCGGGGTGTTTTCAGCATATATGCACTGCTATGCTTGAATTTCCTTGTGTCG 2319
    |||||
Db 252 TTGACTGCGTGAAGTGTGGAATCTACGAGTGTCTGATAGCTTTCCTGTTTG 311
    |||||
QY 2320 TCTTCTTAAGTCCCTGAGACAAGGGGATGCCCTTGAGGTTAATACCGAATTCCTTG 2379
    |||||
Db 312 TCTTCATGAAGTGGCCGAGACAAGGGGATGCCCTTGAAGTCAATCACCGAGTTCTTCT 371
    |||||
QY 2380 CAGTTGTCGCAAGCAAGC 2398
    |||||
Db 372 CTGTCGAGCAAGCAGGC 390
    |||||
```

Search completed: January 5, 2004, 14:13:20
Job time : 746 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 10:29:56 ; Search time 163 Seconds
(without alignments)
7519.760 Million cell updates/sec

Title: US-10-051-909-31
Perfect score: 2777
Sequence: 1 cacgsggttagatcgcgagc.....aaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents NA:*
- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PCtUS_COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200.2	79.2	2824	4	US-09-291-922-1 Sequence 1, Appli
2	940.2	33.9	2601	4	US-09-291-922-7 Sequence 7, Appli
3	691.2	24.9	1487	4	US-09-291-922-13 Sequence 13, Appli
4	594.6	21.4	1692	4	US-09-291-922-9 Sequence 9, Appli
5	394.4	14.2	1009	4	US-09-291-922-15 Sequence 15, Appli
6	227	8.2	870	4	US-09-291-922-5 Sequence 5, Appli
7	143.4	5.2	155	4	US-09-313-294A-4597 Sequence 4597, Ap
8	123.2	4.4	510	4	US-09-291-922-11 Sequence 11, Appli
9	114.2	4.1	443	4	US-09-291-922-3 Sequence 3, Appli
10	96.2	3.5	1853	4	US-09-291-922-23 Sequence 23, Appli
11	84	3.0	2017	4	US-09-291-922-21 Sequence 21, Appli
12	78.2	2.8	2089	4	US-09-291-922-25 Sequence 25, Appli
13	69	2.5	7218	1	US-08-232-463-14 Sequence 14, Appli
14	65.8	2.4	1872	4	US-09-291-922-27 Sequence 27, Appli
15	63.4	2.3	1914	4	US-09-291-922-19 Sequence 19, Appli
16	61.6	2.2	7218	1	US-08-232-463-14 Sequence 14, Appli
17	46.6	1.7	1462	4	US-09-672-749-12 Sequence 12, Appli
18	46.6	1.7	3080	4	US-09-099-041A-25 Sequence 25, Appli
19	46.6	1.7	3080	4	US-09-245-281-25 Sequence 25, Appli
20	46.6	1.7	3080	4	US-09-207-359B-25 Sequence 25, Appli
21	46.6	1.7	3080	4	US-09-340-620A-25 Sequence 25, Appli
22	46.4	1.7	1074	3	US-09-248-335-67 Sequence 67, Appli
23	46.4	1.7	3186	4	US-09-205-258-32 Sequence 32, Appli
24	46.4	1.7	3828	4	US-09-221-013A-7 Sequence 7, Appli
25	46.2	1.7	2136	4	US-09-996-243-302 Sequence 302, App
26	46.2	1.7	441529	3	US-09-103-840A-1 Sequence 1, Appli
27	45.8	1.6	1319	2	US-08-504-459-7 Sequence 7, Appli

28	45.6	1.6	1129	4	US-09-227-357-40	Sequence 40, Appli
29	45.6	1.6	1313	4	US-09-149-476-112	Sequence 112, App
30	45.2	1.6	5852	1	US-07-867-106-2	Sequence 2, Appli
31	44.4	1.6	1447	4	US-09-443-041A-27	Sequence 27, Appli
32	44.4	1.6	1493	1	US-08-340-820-24	Sequence 24, Appli
33	44.4	1.6	1493	1	US-08-593-535-24	Sequence 24, Appli
34	44.2	1.6	7724	4	US-08-486-049-1	Sequence 1, Appli
35	44	1.6	1441	3	US-08-821-994-63	Sequence 63, Appli
36	44	1.6	1736	3	US-09-182-816-22	Sequence 22, Appli
37	44	1.6	1736	3	US-09-182-816-24	Sequence 24, Appli
38	44	1.6	1736	3	US-09-471-528-22	Sequence 22, Appli
39	44	1.6	1736	3	US-09-471-528-24	Sequence 24, Appli
40	44	1.6	1736	3	US-09-634-530-22	Sequence 22, Appli
41	44	1.6	1736	3	US-09-634-530-24	Sequence 24, Appli
42	44	1.6	3581	2	US-08-738-349-1	Sequence 1, Appli
43	43.8	1.6	272	4	US-09-016-434-855	Sequence 855, App
44	43.8	1.6	554	4	US-09-696-169A-14	Sequence 14, Appli
45	43.8	1.6	703	3	US-09-313-300-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-291-922-1
; Sequence 1, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (622)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (636)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (638)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (669)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (771)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (822)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (856)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (889)
; FEATURE:
; NAME/KEY: unsure

LOCATION: (896)
FEATURE:
NAME/KEY: unsure
LOCATION: (944)
US-09-291-922-1

Query Match 79.2%; Score 2200.2; DB 4; Length 2824;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;

Qy	133	CTTCACAGGACCCGGTGACCTCGACGATATCTTGAGACAAAGATGCGGGGCTGTT	192
Db	194	CTTGGCCGCGCTGCCGAGTGGGGCGTAGATTCCGGCGGCATGGGGGGCGCGTGA	253
Qy	193	TTGTGCCATAGTCGCTCCATCGCAATCTATTGACAGGGTGGACAATGCCACCATCG	252
Db	254	TGGTCGCCATCGCGCCTCTATCGCACTTGCTGCAGGCTGGACAATGCGACAATTG	313
Qy	253	CAGCTGCTGTTCTGTATATAAAGAAATTTCAATTGCAAAATGAGCCCACTGTGAGG	312
Db	314	CTGAGCCGCTCCTGTACATAAAGAGAAATTCAACTGCAGAGCGAGCTGTGAGAG	373
Qy	313	GACTAATGTGTCAATGTCATCTTATCGCGCCCAACCATCGTTACTACTATTCTCCGGCCAT	372
Db	374	GCCTCATCGTCCCATGTTCTCTCATTTGGGGCAACATCATCACACATCTCCGGGGCCAA	433
Qy	373	TATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTCTCATTTCTGACTTCTTCA	432
Db	434	GGGCTGACTGCTGTGTAGGAGGCCCATGCTGCTCGCTCGGCTGCTCTACTTCTGCA	493
Qy	433	GCGGCTCATCATGCTATGCTCTCTAATGTCTATGTCCTGCTGTGGACAAGCTTCTGAG	492
Db	494	GTGGCTGTGATGCTTTGGCGCCAAATGTGTACATCTTGTCTCTCGCAAGGCTCATTG	553
Qy	493	ATGGAATTTGGTATTTGGCTTGCTGCTGACGCTTGCTTGACATTTCAGAATAGCCC	552
Db	554	ATGGGTTCGATCGGTTTGCGGTACACTTGTCTCTCATCTCTCGAAACTGCAC	613
Qy	553	C--TTCCGAGATTAGAGGTTTGTGTAATACTACCAATTCAGTG--ATCAGAG	606
Db	614	CGGACAGANATTTTGGGGCTGNTNGAACAGTTCGCCGAGTTCATTGGGGTCAANGAG	673
Qy	607	GAATGTTCTTGTCACTACTGATGTTGGGATGTCCTGTGCGCATCACCCGATTGGA	666
Db	674	GGATGTTCTCTCTACTGATGTTGGGATGTCCTCATGCCCAAACTGATTGGA	733
Qy	667	GAATTATGCTTGGTGTGCTCGGATACCTTCATGTTCTTCTTGGTTGACAATATTTT	726
Db	734	GGGTCACTTGGAGTTCTGTGATCCGTCACCTTATNTACTTTGACTGACTGTCTTCT	793
Qy	727	ATCTTCTGAATCTCCAGATGCTCGTTAGCAAAAGTCCGATGGCAGAGGCAAAAAAG	786
Db	794	ACTTGCCTGAATCACCAAGTGTGTTAGCAAAAGAGATGGCGAGCGAAGAGAG	853
Qy	787	TGTTGCAAAAGTTACGGGGGAAAGACGATGTTCAAGTGAATGTCCTTCTTCTGAAG	846
Db	854	TGNTGCAAGGCTGCGGGGAAGAGATGTTCTCANGGAGANGGCTTCTAGTTGAAG	913
Qy	847	GCTTGAGGTTGGAGAGACACTTCCATGACAGTACATCATTTGACCTGCCACCGAGG	906
Db	914	GTTTGGGGGTCCGTAAGATACACGTAATTNAGAGTACATCATTTGACCTGCCACCGAGG	973
Qy	907	CAGCCGATGATCTTGTACTGACGCGTGAAGAAACAATCACACTTTATGGGCTGAAG	966
Db	974	CAGCCGATGATCTTGTAACTGACGCGTGAAGAAACAATCACACTTTATGGGCTGAAG	1033
Qy	967	AAGGCCAGTCATGATTTGCTCGACCTTCTAAGGGAACCCATCATGCTTGAAGTGTCTTT	1026
Db	1034	AAGGCCAGTCATGATTTGCTCGACCTTCTAAGGGAACCCATCATGCTTGAAGTGTCTTT	1093
Qy	1027	CTCTTGCACTCTGTCATGGAGCATGTTGAACAGAGTACCCCTTATGGATCCGATTG	1086
Db	1094	CTCTTGCACTCTGTCATGGAGCATGTTGAACAGAGTACCCCTTATGGATCCGATTG	1153

Qy	1087	TGACACTTTTGTAGTGTCCATGAGAAATATGCTTCAAGCTGGAGGAAGTATGAGAGCA	1146
Db	1154	TGACACTTTTGTAGTGTCCATGAGAAATATGCTTCAAGCTGGAGGAAGTATGAGAGCA	1213
Qy	1147	CATTGTTTCCAAACTTTGAGATATGTTCAAGTGTCAAGATTCAGATGCCAAAAATGAGC	1206
Db	1214	CATTGTTTCCAAACTTTGAGATATGTTCAAGTGTCAAGATTCAGATGCCAAAAATGAGC	1273
Qy	1207	AGTGGATGAAGAAATCTTCATAGGATGACGAGAGTACGATCTGATGTTGAGAG	1266
Db	1274	AGTGGATGAAGAAATCTTCATAGGATGACGAGAGTACGATCTGATGTTGAGAG	1333
Qy	1267	GTCATATGAGGACATCTCCATAGCCCATTTGCTGTCCAGGCAAGCAAGTCCGGAAG	1326
Db	1334	GTCATATGAGGACATCTCCATAGCCCATTTGCTGTCCAGGCAAGCAAGTCCGGAAG	1393
Qy	1327	GGAAGACATTTGTGACCATGCTTCAACCGTGAAGTGTCTTTGAGCATGAGAAAGCAACC	1386
Db	1394	GGAAGACATTTGTGACCATGCTTCAACCGTGAAGTGTCTTTGAGCATGAGAAAGCAACC	1453
Qy	1387	TCTTAGGGAGGGTGAAGATGCTGAGCAGCAGCATGATATCGGTGGGATGGCAGCTTG	1446
Db	1454	TCTTAGGGAGGGTGAAGATGCTGAGCAGCAGCATGATATCGGTGGGATGGCAGCTTG	1513
Qy	1447	CTTGAAATGTCAGAGAAAGAGTGAAGATGTTAGAAAGAAAGTGTTCAAAAGAG	1506
Db	1514	CTTGAAATGTCAGAGAAAGAGTGAAGATGTTAGAAAGAAAGTGTTCAAAAGAG	1573
Qy	1507	TCTACTTGACCAAGAGGAGTCTCTGCTCAAGAAAGGCTCAATTGTTCACTTCCCG	1566
Db	1574	TCTACTTGACCAAGAGGAGTCTCTGCTCAAGAAAGGCTCAATTGTTCACTTCCCG	1633
Qy	1567	GTCGTGCGGATGTTTGTAGGGTAGTGAGTTGTATCATGCTGCTGCTTATGTAAGTCAGT	1626
Db	1634	GTCGTGCGGATGTTTGTAGGGTAGTGAGTTGTATCATGCTGCTGCTTATGTAAGTCAGT	1693
Qy	1627	CAGCACTTTTCTCAAAGGCTTGTCTGAACCAAGCATGTGAGTCTGCCATGTTTCAAC	1686
Db	1694	CAGCACTTTTCTCAAAGGCTTGTCTGAACCAAGCATGTGAGTCTGCCATGTTTCAAC	1753
Qy	1687	CATCTGAGTAGCTGCAAGGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAAGC	1746
Db	1754	CATCTGAGTAGCTGCAAGGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAAGC	1813
Qy	1747	GTCGCTGTTAGTGGTGTGGAATTCAAGATCTTCAACAGTTTGTCTGGAATAACGCTG	1806
Db	1814	GTCGCTGTTAGTGGTGTGGAATTCAAGATCTTCAACAGTTTGTCTGGAATAACGCTG	1873
Qy	1807	TTCGTACTATACCCCAAAATTTCTGAGCAAGCTGTGTGGCAGTTATTTCTTCAAAAT	1866
Db	1874	TTCGTACTATACCCCAAAATTTCTGAGCAAGCTGTGTGGCAGTTATTTCTTCAAAAT	1933
Qy	1867	TTCGTCTCAGCTCGGATTCAGATCTTGTGATCAAGTTCTCTCACTACCTTACTAATGC	1926
Db	1934	TTCGTCTCAGCTCGGATTCAGATCTTGTGATCAAGTTCTCTCACTACCTTACTAATGC	1993
Qy	1927	TTCCTTGCAATTTGGCTTGGCATGCTCTTATGAGATCTTCCGGAAGAGGTTTGTGCTGC	1986
Db	1994	TTCCTTGCAATTTGGCTTGGCATGCTCTTATGAGATCTTCCGGAAGAGGTTTGTGCTGC	2053
Qy	1987	TAGGACAATTTCCAAATCTGATAGCATCTAGTTATCTGTTGTGCCAATCTAAATTG	2046
Db	2054	TAGGACAATTTCCAAATCTGATAGCATCTAGTTATCTGTTGTGCCAATCTAAATTG	2113
Qy	2047	ATTGGGTACACTAGCCCATGCTTGTCTCCACCGTCAAGTGTATCGTCTACTTCTGCT	2106
Db	2114	ATTGGGTACACTAGCCCATGCTTGTCTCTCCACCATGATGTATCGTCTACTTCTGCT	2173
Qy	2107	GCTTGTATAGGATTTGTTGCCATTCGCCAACAATTTATGTGACAGAGATCTTCCAAACA	2166
Db	2174	GCTTGTATAGGATTTGTTGCCATTCGCCAACAATTTATGTGACAGAGATCTTCCAAACA	2233

QY 2167 GGGTTCGTCCTCTGTATGTCATTTGTCCTTACATTTCTGATCGAGATATCATCG 2226
 Db 2234 GGGTTCGTCCTCTGTATGTCATTTGTCCTTACATTTCTGATCGAGATATCATCG 2293
 QY 2227 TCACCTACAGCCTTCTGTGATGCTGAATGCTATTGAGCTGGCGGTGTTTCAGCATAT 2286
 Db 2294 TCACCTACAGCCTTCTGTGATGCTGAATGCTATTGAGCTGGCGGTGTTTCAGCATAT 2353
 QY 2287 ATGAGTCGTATGCTGATTTCTTTGTCCTTCTGTCCTTAAAGTCCCTGAGCAAAAG 2346
 Db 2354 ATGAGTCGTATGCTGATTTCTTTGTCCTTCTGTCCTTAAAGTCCCTGAGCAAAAG 2413
 QY 2347 GGATGCCCTTGAGGTATATACCAATTCTTGCAAGTTGGTGGCAAGCAGCGCTGCA 2406
 Db 2414 GGATGCCCTTGAGGTATATACCAATTCTTGCAAGTTGGTGGCAAGCAGCGCTGCA 2473
 QY 2407 AAGCCTAATTTCTTTGTCCTTTGTCCTTCTGTCCTTAAAGTCCCTGAGCAAAAG 2466
 Db 2474 AAGCCTAATTTCTTTGTCCTTTGTCCTTCTGTCCTTAAAGTCCCTGAGCAAAAG 2533
 QY 2467 GGGGTTTCAACCAAGAGCTCGAGAAATTAATTTGATTTGTTAAAGGAAAGCA 2526
 Db 2534 GGGGTTTCAACCAAGAGCTCGAGAAATTAATTTGATTTGTTAAAGGAAAGCA 2593
 QY 2527 ACATCTGCTCATGCTCCTCAACCGGTAAAGAGTCCCTCAATGGCAATAGAGTCTG 2586
 Db 2594 ACATCTGCTCATGCTCCTCAACCGGTAAAGAGTCCCTCAATGGCAATAGAGTCTG 2653
 QY 2587 TAAGTTCATATGTCATTTACCATATGTTTAACTATTTGATTTAATTAAGTCAAG 2646
 Db 2654 TAAGTTCATATGTCATTTACCATATGTTTAACTATTTGATTTAATTAAGTCAAG 2713
 QY 2647 TATTCACGCTGCTGTTGTCGTAGAAATCTTTAGAACAAAGATGATTAATGATCTGA 2706
 Db 2714 TATTCACGCTGCTGTTGTCGTAGAAATCTTTAGAACAAAGATGATTAATGATCTGA 2773
 QY 2707 TGTATTAATTAATCAAAATCTCAAAATTAAGAAATATCGTTTCTCAAAAAA 2757
 Db 2774 TGTATTAATTAATCAAAATCTCAAAATTAAGAAATATCGTTTCTCAAAAAA 2824

RESULT 2

US-09-291-922-7
 : Sequence 7, Application US/09291922
 : Patent No. 6383776
 : GENERAL INFORMATION:
 : APPLICANT: Allen, Steve
 : APPLICANT: Hitz, Bill
 : APPLICANT: Kinney, Tony
 : APPLICANT: Tingey, Scott
 : TITLE OF INVENTION: Plant Sugar Transport Proteins
 : FILE REFERENCE: BB-1163
 : CURRENT APPLICATION NUMBER: US/09/291,922
 : EARLIER FILING DATE: 1999-04-14
 : EARLIER APPLICATION NUMBER: 60/083,044
 : NUMBER OF SEQ ID NOS: 30
 : SOFTWARE: Microsoft Office 97
 : SEQ ID NO 7
 : LENGTH: 2601
 : TYPE: DNA
 : ORGANISM: Glycine max
 US-09-291-922-7

Query Match 33.9%; Score 940.2; DB 4; Length 2601;
 Best Local Similarity 66.0%; Pred. No. 1e-241;
 Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

QY 177 ATGCGGGGCTGTTCTTGTGCGCATAGTCCTCCATCGGCAATCTAFTGCAGGGGTG 236
 Db 175 ATGAAAGTGCCTGCTTGTGCTATGCGCTTCATGTAATTTCTCAAGGATG 234
 QY 237 GACAAATGCCAATCGCAGCTGCTGTTCTGTATATAAAGAAATTCAATTGCAAAAT 296

Db 235 GATAATGCTACCATCGCGGGGCTAATGTTACATTAAGAAAGACCTTGCTTTGGAA-- 292
 QY 297 GAGCCCACTGTGAGGACTAATGTGTCAATGTCATTTATCGCGCCACCATCGTTACT 356
 Db 293 ----CAACTATGAAAGGCTTGTGTGGGATGTCCTGATTTGAGCAACGGTATACAC 348
 QY 357 ACATTTCTCGGGCCATTAATCAGACTCGATTGGCCGAGCCCTATGCTTATTTCTTTCA 416
 Db 349 ACATGCTCTGTCTTATAGCGGATTTGCTCGGTCCGCAACCATGATGATTAATCTCAT 408
 QY 417 ATTCTGTAATTTCTTACGCGCTCATCATGCTTAAGTCTCTTAATGTCATGTCCTG 476
 Db 409 GTGCTTATTTCTTGGGTGTTGTGATGCTGTGTCCTCAATGTATGTGTGTC 468
 QY 477 TTGGCAGCTTCTAGATGATTTGGTATTTGGCTTGGCTGTCACGCTTGCTTGTAC 536
 Db 469 TTGGCAGGCTACTTGATGATTTGGGATTTGGCTTGGCTGTCATCTTGTCCGCTAT 528
 QY 537 ATTTCAAAATAGCCCTTCGAGATTAAGGTTTGTGAATACACTACCAATTCAGT 596
 Db 529 ATATCGAAACGGCCGCTTGAATTAAGGGGTCGTTGAATAACGCTTCTCAGTTCACT 588
 QY 597 GATCAGGAGGAATGTTCTGTCTATCTGATGCTGTTGGGATGTCCTGTCGCACTCA 656
 Db 589 GGTCTGAGGAATGTTTGTCTACTGTATGTTTGGCATGTCATGATGCCGCG 648
 QY 657 CCCGATTGAGAAATTAATGCTGTGTGCTCGGATCTTCAATGTTCTTGTGTTG 716
 Db 649 CTTAGCTGAGGCTCATGCTGGGCTTGTCTATCTCTCTCTGATTTTGCATG 708
 QY 717 ACAATATTTTATCTTCTGTAATCTCAAGATGCTGTTAGCAAAAGTGGATGGCAG 776
 Db 709 ACAATTTTCTTCTGCGGAGTCTCCTCGGTGCTGTCAGCAAGAAAGATGCTGAG 768
 QY 777 GCAAAAAGGTGTGCAAAAGTTACGGGGGAAAGACATGCTCAGGTGAATGTCCTT 836
 Db 769 GCTAAGAAAGGTGCTCAAAAGATTCGCGGAAAGGAGATGTCTCAGGATGGCAT 828
 QY 837 CTTCCTGAAGGTTGAGGTTGAGGAGACATTTCAATGAAGATGATCATTTGACCT 896
 Db 829 CTGCTGAAGGCTCTCGGATTTGGGGGTGATATCATCTATCGAAGATCATTAATTTGCC 888
 QY 897 GCCACGAGGAGCCGATGATCTTGTACTGACGGTGAAGAAACAATCACACTTTAT 956
 Db 889 GCTGACGATGTGCTGATGTCATGAACATGCAACAGAAAGATTAATTCATTAAT 948
 QY 957 GGGCTGAAGAGGCAATCATGATTTGCTGACCTTCTAAGGAGCCCATCATGCTTGA 1016
 Db 949 GATCCCAAGCAGGCTTCTTGTGTTATCAAAACCTGTCACTGACAGAGTTCTATTGG- 1007
 QY 1017 AGTGTCTTCTCTGATCTGTCATGAGGAGCATGGAACCAAGATGACCCCTTATG 1076
 Db 1008 -----CCTGCGTACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
 QY 1077 GATCCGATTTGACACATTTTGTGATGTCATGAGATATGCTCA-----AGCTGA 1130
 Db 1057 GATCTCTGTGACACATGTTGTGATGATTCATGAGAACTCCCGAGACAGAGCAAGA 1116
 QY 1131 GGAATATGAGGACACATGTTTCCAACTTTGGAAGTATGTTCAAGTCAAGATCAG 1190
 Db 1117 GGAAGCATGCGAAGCAGCTGTTCCAAATTTGGAAGATGTTCAAGCATGCTGAGCG 1176
 QY 1191 CATGCCAAAATGAGCAGTGGATGAAGAAATCTTCAATAGGATGACGAGATGACGA 1250
 Db 1177 CATGCTAAATTTGAACAATGGATGAAGAAAGCTTACAAAGGAACTGAGGATCATG 1236
 QY 1251 TCTGATGTGACGAGGTGATATGAGCAATCTCCATAGCCCATTTGCTCCAGCAG 1310
 Db 1237 TCAGATCAACCCGTGGGACTCCGATGATTAATTTGCAAGTCTTTAATCTCACGCCAA 1296
 QY 1311 GCACAGGTGCGAAGGAAAGACATTTGTCACCATGCTCAACCTGGAAGTCTTTGAGC 1370

```
Db 1297 ACAACAGCCTTGAAAAAGACTTACCTCCTCCTCCTTCCCATGGCAGTATCCTTGGCAGC 1356
QY 1371 ATGAGAGGCA---AACCTCTTAGGGGAGGGTGGAGATGCTGTGACGACACTGATATC 1427
Db 1357 ATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAAGGTGAGCAAGGTGTAGTACAGGTATT 1416
QY 1428 GGTGGGGGATGGCAGCTTGCTTGGAAATGGTCAAGAAAGGAAGTGAATGTGTAGAAAG 1487
Db 1417 GGTGGTGGCTGGCAACTGGCATGGAATGGACTGATTA--AGTGAGGATGAAAAACAA 1473
QY 1488 GAAGGTGGTTTCAAAAAGAGTCTACTTGCAACCAAGAGGAGTTCTGGCTCAAGAGGGGC 1547
Db 1474 CAAGGAGGGTTTAAAAAGATTATTTACATGAGGAGGAGTTTTCATCTCGTGTGA 1533
QY 1548 TCAATTGTTTCACTTCCCGGTGGCGAGTGTTTTGAGGGTAGTGAAGTTGTACATGCT 1607
Db 1534 TCCATGTATCGATTCCCGGTGAAGGCGAA-----TTGTCCAGGCT 1575
QY 1608 GCTGCTTAGTAAGTCAAGTCAAGCACTTTCTCAAGGCTCTTGCTGAACCAAGCATGTCA 1667
Db 1576 GCTGCTTGTGTAAGCCCAACCCGCTCTTACTCCAAGAGCTTATGATGACACCCAGTT 1635
QY 1668 GATGCTCCATGGTTCACCCATCTGAGGTAGCTGCCAAAGGTTCAAGTTGGAAGATTGG 1727
Db 1636 GGGCCTTGCATGGTTTACCCCATCTGAGACAGCTTCAAGGGGCCCAAGTTGAAAAGCTCTT 1695
QY 1728 TTGGAACCTGAGTGAGGCGGTGCTGTTAGTGGGTGTTGAATTCAGATCCTTCAACAG 1787
Db 1696 CTGGAACAGGGGTTAAGCATGATGGTGTGGNGTTGGAATACAAATACTTCAGCAG 1755
QY 1788 TTGCTGGAATAAAAGGTTCTGTACTATACCCCAAAATCTTGAGCAAGCTGTGTG 1847
Db 1756 TTTTCAGGGAATAATGGGTTCTATATTACACCTCAAAATCTTGAAGAGCGCGTGT 1815
QY 1848 GCAGTTATCTTTCCAAATTTGCTGCTGAGCTGGCATCAGCATCCATCTTGATCAGTTCT 1907
Db 1816 GAAGTCTCTTTCAGATATAGCATGGCTCAGAGTCGSCATCAATCCTTATCAGTGTCT 1875
QY 1908 CTGACTACCTTAATAATGCTTCCCTGCAATGGCTTTGCCATGCTGCTTATGGATCTTCC 1967
Db 1876 TTGACAACTTCTTGATGCTTCCCTGTATAGGCGGTAGCCATGAAGCTCATGGATGTTCA 1935
QY 1968 GGAAGAGGTTTTTGTGCTAGGCAAAATTCCAATCTTGATAGCATCTAGTATCTGTG 2027
Db 1936 GGGAGAGGCAAGTGTACTTACTTACAAATCCCGTGTGATGTGTCATCTATATTG 1995
QY 2028 GTGTGTCCAATCTAATGATTGGGTACACTAGCCCATGCTTGTCTCCACCGTCACT 2087
Db 1996 GTGATTGGAAGCCTGTAAATTTGGCAATGTGCCCATGAGCAATCTCAACAGTATGC 2055
QY 2088 GTTATGCTACTTCTGCTGCTGTATGGGATTGGTCCCATCCCAACATTTATGT 2147
Db 2056 GTTGTGTTATTTCTGCTGCTTGTGATGGGTTATGAGCAAAATCCAAACATCCTTGC 2115
QY 2148 GCAGAGATCTTCCAAACAGGTTCCGTGGCTCTGTATGGCAATTTGTGCTTACATTC 2207
Db 2116 TCAGAGATTTTCCCACTAGGGTGGCTCTGTCAATGCTATCTGTGATTAGTGTTC 2175
QY 2208 TGGATCGAGATATCATGCTCAACCTTACAGCTTCTGTGATGCTGAATGCTATTGGACTG 2267
Db 2176 TGGATTGAGACATCATCATCACTATCTGCTGCTGCTGTGATGCTGGCTCTTTAGGACTT 2235
QY 2268 GCGGGTGTTTTACGATATATGCAAGTGTATGCTTATTTCTTGTGTTGCTTCTT 2327
Db 2236 GGTGTGTATTCGCCATTACGCAAGTGTGTGTTTCATCTGTGATATTTGTGTTTGG 2295
QY 2328 AAGTCCCTGAGCAAAAGGGGATGCCCTTGAGGTATATTACCGAATCTTTGAGTTGGT 2387
Db 2296 AAGTTCCAGAAACAAAGGCGATGCCCTTGAAGTCACTCTGAATCTTTTCTGTGGA 2355
QY 2388 GCGAAGCAAGCGGCTGC 2404
Db 2356 GCAAGGAGGCTGCTTC 2372
```

```
RESULT 3
US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-13

Query Match      24.9%; Score 691.2; DB 4; Length 1487;
Best Local Similarity 80.1%; Pred. No. 3.6e-175;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGTGAGATGCTGTGACAGCACTGATATCGGTGGGGATGGCAGCTTGTGAAA 1454
Db 12 GAGGTTGGGAGGAGTCAGCACTGTGATTTGGTGGGGGTGCAACTCGCATGAAA 71
QY 1455 TGGTCAGAGAAAGGATGAGATGTGAGAAAGAGAGGTGTTTCAAAAGAGTCTACTG 1514
Db 72 TGGTCGAGCGACAGGCGAGATGGCAAGAGAGAGGCTTCAAAAGATCTACTTG 131
QY 1515 CACCAAGAGGAGTTCCTGCTCAAGAGGGGCTCAATTGTTCACTTCCCGTGTGGC 1574
Db 132 CACCAAGAGGAGTGGCCGACTCAAGAGGGGCTGTGTTTCACTTCTGTGGGGGT 191
QY 1575 GATG---TTTTAGGGTAGTGAAGTTGTGACATGCTGCTTGTAGTACATGACGA 1631
Db 192 GATGCCACGCAAGGGGAGTGGGTTTATACATGCTGCTTGTGTAAGCACATCGGCT 251
QY 1632 CTTTCTCAAGGCTCTGCTGAACCAAGCATGTGAGTGT---GCCATGTTCAACCA 1688
Db 252 CTTTACTCCAAGATCTTATGGAAGAGCGTATGCGCGCCGTCACGATGATTCACA 311
QY 1689 TCTGAGTAGTGTGCCAAAGTTCAAGTTGGAAGATTTGTTGAACCTGAGTGAAGCT 1748
Db 312 TTGAGGCGAGCTCCCAAGGTTCAATCTGGAAGATCTGTTGAACCTGCTGAGGCT 371
QY 1749 GCCCTGTAGTGGGTGTTGAATTCAGATCTTCAACAGTTTGTGGAATAAACGTTG 1808
Db 372 GCATGTGTCGTGGTGTGGAATTCAGATGCTTCAAGAGTTGCTGGAATAATGAGTT 431
QY 1809 CTGTAATATACCCCAAAATTTTGAGCAAGCTGTGTGGCAGTTATTTCTTCCAAATT 1868
Db 432 CTCTACTATACTCTCAAAATTTGAGCAAGCTGTGTGGCTGTCTTCTTCCAAATCTT 491
QY 1869 GGTCTCAGCTGGCATGACATCCATCTGATCAGTCTCTCACTACCTTACTAATGCTT 1928
Db 492 GGCTCAGTTACGATCAGATCCATCTGATCAGTCTCTCACCACTTACTCATGCTC 551
QY 1929 CTTGCAATGGCTTGGCATGCTGCTTATGATGATCTTCCGAAGAAGGTTTGTGCTA 1988
Db 552 CCAAGCATTTGTTGATGAGTGAAGTATGATGATATCTGGAAGAAGGTTTCTGCTA 611
QY 1989 GGCACAATTCATCTGATAGCATCTCTAGTATCTGTTGTGTCCAATCTAATGAT 2048
Db 612 GGCACAAATTCATCTGATAGCATCCCTAATGTTTGGGTGTGTCCAATGATCAAC 671
```


Db 1298 GCTCAATCTGTAGGCCCTCGCTGCTGTTTGGTATTATGCTGTCGTGTGCTTCATAGC 1357

Qy 2309 CTTTGTGTTGTCCTTCCCTTAAGSTCCCTGAGACAAAGGGGATGCCCCCTTGAGTTATTAC 2368

Db 1358 ATGGGTGTTGTCTTTTGAAGTTCCAGAAACCAAGGGCATGCCACTGGAAGTATCAT 1417

Qy 2369 CGAATTCTTGCAGTTGGTGCGAACA 2395

Db 1418 TGAAGTCTTCTCTGTCCGAGCAAAACA 1444

RESULT 5

```

US-09-291-922-15
; Sequence 15, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kihney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-15

```

Query Match	14.2%;	Score 394.4;	DB 4;	Length 1009;
Best Local Similarity	74.4%;	Pred. No. 8.6e-96;		
Matches 497;	Conservative	0;	Mismatches 171;	Indels 0;
				Gaps 0;

OY		1730	TGAACCTGGAGTGAGGCGGTGCCCTGTTAGTCGGTGTGGAATTCAGATCCTTCAACAGTT	1789
Db		1	TGAACCTGGAGTGAAGCATGCACTGTTGCTGGCATAGATTACAAGATCCGCAGAGTT	60
OY		1790	TGCTGGAATAAACGGGTGTTCTGTACTATACCCCACMAATTCTGAGCAAGCTGTGTGC	1849
Db		61	TGCGGGTATCAATGAGAGTCCCTACTACACACCTCAGATACTTGAGCAAGCAGTGTGCG	120
OY		1850	AGTTATCTTCCCAATTTGGTCTCAGCTCGGCATCAGATCCATCTTGATCAGTTCTCT	1909
Db		121	G GTTCTTCTATCAAACATTTGGACTAAGCTCTTCCCTCAGCATCTATTCTTATTAGTGCCTT	180
OY		1910	C ACTACCTTAATAATGCTTCTTGCATTGGCTTTGCCATGCTGCTTATGATCTTTCGG	1969
Db		181	G A C A A C C T T G C T G A T G C T T C C A G A T T G G C A T G C C A T G A G A C T C A T G A T A T G T C A G G	240
OY		1970	AAGAAGTTTTTGTGCTGCTAGGCAACAATCCAATCTGTATAGCATCTCTAGTTATCCTGCT	2029
Db		241	AAGAAGGTTTCTTCTCTCTTCAACAATCCCTGTCTGTATAGTAGCGCTAGCTGTCTTGGT	300
OY		2030	TGTGTCCAATCTAATTGATTTGGGTACACTAGCCCATGCTTGTCTCCACCCTCAGTGT	2089
Db		301	TTTAGTGAATGTTCTGAGATGTGGGAACCATGCTGACCGCTCGGCTCTCAACGATCAGCGT	360
OY		2090	TATGCTTACTTCTGCTGCTTCGTATGGGATTGTGCCCATCCCAACATTATATGTGC	2149
Db		361	CATGCTTATTTCTGCTTCTTCGTATGGGGTTGGGCTCTATCCCAATATTCTGTGCGC	420
OY		2150	AGAGATCTTCCAACACAGGGTTCGGGCTCTGTATTGSCAATTGTGCCTTACATTCTG	2209
Db		421	GGAGATTTTCCCAACCTCTGTGCGGTGCATCTGCATAGCCATCTGCGCGCTAACCTTCTG	480
OY		2210	GATCGAGATATCATCTGTCACTCACAGCCTTCTCTGTGATGCTGAATGCTATTGGACTGGC	2269
Db		481	GATCGCGACATCATCTGTGACATACACTCTCCCGGTAGTCTCAATGCCAATTGGTCTCGC	540

OY	2270 GGGTGTTCAGCATATATGCAGTCGTAAGCTTGATTCCTTGTGTGTCGTCTTCTTAA	2329
Dd	541 TGAAGTCTTCGGCATATATGCCATCGTTGTGTAACAGCCTTGTATTGTTACATGAA	600
OY	2330 GGTCCTGAGACAAGGGGATGCCCTTGAGGTATAACGAATCTTTGCAGTTGGTGC	2389
Dd	601 GGTCCCTGAGACAAGGGCATGCCCTGAGGTCATCACCGAAGTCTCTCTGTGCGGGC	660
OY	2390 GAAGCAAG	2397
Dd	661 AAAGCAGG	668

RESULT 6

```

; Sequence 5, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-5

```

Query Match	8.2%	Score 227;	DB 4;	Length 870;
Best Local Similarity	74.9%;	Pred. No. 5.3e-51;		
Matches 284; Conservative	0;	Mismatches 95;	Indels 0;	Gaps 0;

QY	2020	TTATCCTGGTTGTGTCCAACTTAATTGATTTGGGTACACTAGCCCATGCTTGTCTTCCA	20739
Db	12	TAACTTGATTCGTGTCAATATTTCTGGATGTGGGACCATGGTTCAATGCTCACTGTCCA	71
QY	2080	CCGTCAGTGTATTCGTCTACTTCTGTGCTTCGTTATGGGATTTGTGCCATCCCCAACA	21339
Db	72	CAGTCAGTGTCACTACTTACTTCTGTCTTCTTTGTCAATGGGGTTCGGGCTTATTCAAACA	131
QY	2140	TTTTATGTGACAGATCTTTCCAAACGAGGTTCCGTGCCCTCTGTATTGCCATTTGTGCT	21999
Db	132	TTCTCTGTGACAGATTTTCCGACCACCGTTCGTGCATCTGCATAGCCATCTGTGCC	191
QY	2200	TTACATTCTGGATCGGAGATATCATCGTCACTACGCTTCTGTGATGCTGAATGCTA	22599
Db	192	TAACTTCTGGATCGGTGATATCATTTGTGACATACACCTCCCGTGATGCTCAACGCCA	251
QY	2260	TTGGACTGCGGGGTGTTTTCAGCATATATGACAGTGTGTAATGCTTGATTTCTTTGTGTTG	23199
Db	252	TTGGACTCGCTGGAGTGTTTGGAAATCTACGCAGTGTGTCTGATACTGGCTTTCCTGTTTG	311
QY	2320	TCTTCTTAAGTCCCTTGAGACAAAGGGGATGCCCTTGAAGTTATTACCGAATCTTTTG	23799
Db	312	TCTTCATGAAGGTCCCGAGACAAAGGGCATGCTCTTGAAGTCATCACCGAGTTCTTCT	371
QY	2380	CAGTTGGTCCGAAGCAAGC	2398
Db	372	CTGTGAGCAAGCAGGC	390

RESULT 7

US-09-313-294A-4597
; Sequence 4597, Application US/09313294A

```
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4597
; LENGTH: 155
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348695H1
; NAME/KEY: unsure
; LOCATION: 45, 56, 77, 83, 140, 154
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4597
```

```
Query Match          5.2%; Score 143.4; DB 4; Length 155;
Best Local Similarity 96.0%; Pred. No. 5.5e-29;
Matches 144; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1039 GTCATGGAGCATGTGTGAACCCAGAGTGTACCCCTTATGATCCGATTGTGACACTTTTG 1098
DB 1 GTCATGGAGCATGTGTGAACCCAGAGTGTACCCCTTATGATCCGATTGTGACACTTTTG 60
OY 1099 GTAGTGCATGAGATATATGCTCAAGCTGAGGAGGAGATGAGGACACATTGTTCCAA 1158
DB 61 GTAGTGCATGAGAAATATGCCNCAAGCTGAGGAGGAGATGAGGACACATTGTTCCAA 120
OY 1159 ACTTTGAAGTATGTTCAAGTGTACAGATC 1188
DB 121 ACTTTGAAGTATGTTCAAGTGTACAGATC 150
```

RESULT 8

```
US-09-291-922-11
; Sequence 11, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (421)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (434)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (441)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (458)
```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (483)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (493)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (498)
US-09-291-922-11
```

```
Query Match          4.4%; Score 123.2; DB 4; Length 510;
Best Local Similarity 72.7%; Pred. No. 2.5e-23;
Matches 184; Conservative 0; Mismatches 66; Indels 3; Gaps 2;
```

```
OY 168 GAGACAAGATGTCGGGGCTGTTCTTGTGCGCAATAGTCGCTCCATCGCAATCTATTG 227
DB 194 GGAAGAAGATGTCGGGTGCTGCACTGTGCGCAATGCGGCTTCCATTGGCAATCTGCTG 253
OY 228 CAGGGGTGGACAATGCCACCATCGCAGCTGCTGTTCTGTATATAAGAAGAAATTTCAA 287
DB 254 CAGGGGTGGACAATGCCACCATCGCTGCTGCTGTTCTGTATACAGAAGAAGAAATTTCCAG 313
OY 288 TTGCAAAATGAGCCCACTGTGAGGGAGCTAATTGTGCAATGTCATCTATCGGCGCAC 347
DB 314 CTGAAAATATATCCGACTGTGAGGGGCTCATGTCGCA--TGTCTCATCGGGTGCAC 371
OY 348 ATCGTACTACATTCCTCGGGCCATTATCAGACTGATGCGCGACGCCCTATGCTATT 407
DB 372 ATCATCAC-ACATTCCTCGGGCCAGTATCAAACTGGGTGCGGGCTTANGCCATCTCC 430
OY 408 CTCTCTCAATTC 420
DB 431 TTGTTTCAATTC 443
```

RESULT 9

```
US-09-291-922-3
; Sequence 3, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (193)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (388)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (435)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (439)
US-09-291-922-3
```

```
Query Match          4.1%; Score 114.2; DB 4; Length 443;
Best Local Similarity 74.5%; Pred. No. 6.1e-21;
```

	Matches	143;	Conservative	0;	Mismatches	49;	Indels	0;	Gaps	0;
Oy	177	ATGTCGGGGGCTGTTCTTGTCCCATAGTCGCGCTCCATCGCAATCTATTGCAGGGGTGG								236
Db	205	ATGCGCGGCGCGCTGTGTGTCGCATCGCGGCGCTCCATCGCACTGTCTGCAGGGCTGG								264
Oy	237	GACAAATGCCACCATCGCAGCTGCTGTTCTGTATATAAAGAAAGAAATTCAATTGCAGAAAT								296
Db	265	GATTAATGCAACCATTCGAGGTGCGTACTGTACATCAAGAGAAATTCAACTTGCAATGC								324
Oy	297	GAGCCCACTGTGAGGAGACTAATTGTGTCATATGTCACCTTATCGCGCCACCATCGTTACT								356
Db	325	GACCCCTTATCGAAGGTCTGATCTGGCCCATGTGCTCATTTGGGGCCACCATCATCAG								384
Oy	357	ACAATTCTCCGGG								368
Db	385	ACGNTCTCTGCG								396

```

RESULT 10
US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23

```

Query Match	3.5%;	Score 96.2;	DB 4;	Length 1853;
Best Local Similarity	48.6%;	Pred. No. 8.1e-16;		
Matches 297;	Conservative	0;	Mismatches 308;	Indels 6;
				Gaps 1;

[illegible]

QY 606 GGAATGTTCTTGTGCATACTGCATGGTGTTGGGATGTCCTGTGCCCATCACCCGATTGG 665
| | | | |
Db 671 GGGAATTAATTGGATACATATCAAACTATGCATTTTCGAAGCTGACACTAAGTGGGA 730

QY 666 AGAATTATGCTTGGTGTGCTCGCGAATACCCTTCATTGTTCTTGGTTGACAATATT 725
| | | | |
Db 731 TGGCGAATGATGCTTGGAGTTGGTGCAATACTTCGGTACTCCTAACAGTAGAGTGTG 790

QY 726 TAATCTTCCGTGAATCTCCAAGATGGCTCGTTAGCAAAGSTCGGATGSCAGAAGCAAAAAAG 785
| | | | |
Db 791 GCCGATGCCGAGTCCCCCAAGSTGGCTTGTGATGAGGGGTCGTTTGGAGAGGCAAGAAAA 850

QY 786 GTGTGCAAAA 796
| | | | |
Db 851 GTGCTTAACAA 861

```

RESULT 11
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-21

```

Query Match	3.0%;	Score 84;	DB 4;	Length 2017;
Best Local Similarity	47.1%;	Pred. No. 1.6e-12;		
Matches 323; Conservative	1;	Mismatches 356;	Indels 6;	Gaps 2;

Dp	82	CTGCACCA	CCGAGATGG	CTTCGCCGCGCTGC	GAGGCGGTGCGCGGAAGAAGG	141
Qy	172	ACAAGATG	TGGGGCTGTCTT	GTGCGCCATAGTGC	CCTCATTGGCAATCTATTGCAGG	231
Dp	142	GCAACGTCCGGTT	CGCTTGCCCTTGCCG	CCATCCTCGCTTCATGACC	TCCATCCTCCTCG	201
Qy	232	GGTGGACA	TGCCACCATCGCAG	CTGCTGTTCTGTATA	TAAAGAAGAAATTTCAATTGC	291
Dp	202	GCTACGATATCGGG	TGATGACGGGGCGT	CGCTGTACATCAAGAAGACTTCAACAT--		259
Qy	292	AAATGAGCCCA	CTGTGAGGGACTAATT	GTGTAATGCACTTATCGGCGCACCATCG		351
Dp	260	-CAGTAGCCGGAAG	STGAGTTCTCATGGGCATAC	TGAACC--TCTACTCGCTCATCG		315
Qy	352	TTACTACATTCT	CCCCGCCATTATCAGACTCGATTGGCCGACGCC	CTATGCTTATTCTCT		411
Dp	316	GCTCCTTCGGCGGG	CGGCGGACGTCGGACTGATCGGCGCGGTACAC	CAACATCGTGTTCG		375
Qy	412	CTTCAATTCTGT	ACTTCTTTCAGCGGCTCATFCATGCTATG	GTCTCTCTAATGCTATGTC		471
Dp	376	CCGCCGTCA	TATTCTTCGGGGGGSGTTCTCTCATG	GGGTTCCCGTCAACTACGCCATGC		435
Qy	472	TGCTGTTGGCAG	CGTTTCGTAGATG	GATTGGTATTGGCTTGCGCTGCACGCTGTGCTT		531
Dp	436	TCATGTTCCGCC	CGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGGCGCG			495


```
QY 532 TGTACATTTCAGAAATAGCCCCCTTCGGAGATTAGAGTTTGTCTGAATACACTACCACAT 591
    |||||
Db 496 TGTACACCCCGGAGGTGTCCCGCGCTCGCGCGCTGTGCTTCTGACGTCGTTCCCGGAGG 555
QY 592 TCAGTGGATCAGGAGAAATGTTCTGTCTCATCTGCATGGTGTGGAGTGTCCCTGTGCG 651
    |||||
Db 556 TGTTCATCAACTTCGGCATCTGTCTGGGTACGTCCTGAACATATGCTTCTCCCGCTTGC 615
QY 652 CATCACCCGATTGAGAAATATGCTGTGTGCTCGCATACCTTCATTTCTTCTTTG 711
    |||||
Db 616 CGCTGAACCTCGGGTGCGCATCATGCTCGGCATCGGCGCGCGCGCTCCGTGCTGCTCG 675
QY 712 GTTGAACATATTTATCTTCTGTAATCTCCAAGATGGCTCGTTAGCAAGTCGATGG 771
    |||||
Db 676 CGCTCATGTGCTCGGCATCGCGGAGTCCCGCGGTGGCTGTGTCATGAAGGACCGCTCG 735
QY 772 CAGAGGCAAAAAAGGTGTGCAAAAG 797
    |||||
Db 736 CGGACGCCCAAGGTGTGCTGGAGAAG 761
```

RESULT 12

```
US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Linney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
; US-09-291-922-25
```

Query Match 2.8%; Score 78.2; DB 4; Length 2089;
Best Local Similarity 48.2%; Pred. No. 5.7e-11;
Matches 317; Conservative 0; Mismatches 328; Indels 12; Gaps 3;

```
QY 144 CCGGTACCTCGACGATATCTTGAGGACAAGATGTCCGGGCTGTCTTGTGCCATA 203
    |||||
Db 95 CCGGCGCGCTCGAGCCCAAGAGAGGCAACGTGAGTTCCGCTTCGCTGCGCATC 154
QY 204 GTCCCTCCATCGGCAATCTATTGAGGGGTGGACAATGCCACCATCGACGTGCTT 263
    |||||
Db 155 CTCGCTCCATGACCTCATCTCTCGGCTACGACATCGGCGTGATGAGCGAGCGTCG 214
QY 264 CTGTATATAAGAAGAAATTTCAATTCGAAATGAGCCCACTGTGAGGAGTAATTGTG 323
    |||||
Db 215 CTGTACATCCAGAAG-----ATCTGAAGATCAACGACACCACTGAGGTCTCATG 268
QY 324 TCAATGTCACTTATCGCGCCACCATCGTTACTACATTTCCGGGCCATTATCAGACTCG 383
    |||||
Db 269 GGCATCTCAACGTGTACTCGCTCATTTGGCTCTTCGCGGCGGGCGGACGTCCGACTGG 328
QY 384 ATGGCCGACGCCCTATGCTTATCTCTCTTCAATTTCTTACTTCTTACGCGGCTCATC 443
    |||||
Db 329 ATCGGCGGCGCTTCAACCATGCTTTCGCGCGCTCATCTTCTTCGCGGCGCCCTCATC 388
QY 444 ATGCTATGCTCTCTAATGTCTAATGCTGCTGTGTCACGCTTCTAGATGATTTGCT 503
    |||||
Db 389 ATGGCTTCTCTCTCAACTACGCGCATGCTCATGTTGCGGCGCTTCGTGCGCGCATCGC 448
QY 504 ATTGGCTTGTCTACGCTGTGTGCTTTGTACATTTCAGAATAGCCCTTCGAGATTT 563
```

```
Db 449 GTGGGGTACGCTCTCATGATCGCGCCCGGTGAACAGCGGCGAGGTGTCCCCCGCTGCC 508
QY 564 AGAGTTTGTGTAATACACTACCAATTCAGTGAATCAGAGGAATGTTCTTGTATAC 623
    |||||
Db 509 CGTGGGGTCTCAACATCTTCCCGGAGGTGTTTCATCAACTTCGGCATCTCTCGATAT 568
QY 624 TGCATGCTTTGGGATGTC---CCTGTGCCATCACCCGATTGAGAAATATGCTTGT 680
    |||||
Db 569 GTCTCAACTTCGCTTGCCTCGCCGCTCTCCCTCGGCTGCGCATATATGCTCGGC 628
QY 681 GTGCTCGGATACCTTCATGTTCTTCTTTGTTGACAAATATTTATCTTCTGAATCT 740
    |||||
Db 629 ATAGCGCGGTGCCGTC---CGTCTGTCTCGGTTCAATGCTGCTCGCATGCCGAGTCT 685
QY 741 CCAAGATGCTCGTTAGCAAGGTGCGATGCGAGGCAAAAAAGGTGTGCAAAAG 797
    |||||
Db 686 CCGCGGTGCTCGTCATGAAGGCGCGTCTCGCGGACGCCAAGTTGTGCTTGCAGAG 742
```

RESULT 13

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT29pl-F18
; US-08-232-463-14
```

Query Match 2.5%; Score 69; DB 1; Length 7218;
Best Local Similarity 5.3%; Pred. No. 3e-08;
Matches 21; Conservative 227; Mismatches 147; Indels 0; Gaps 0;

[illegible]

```

RESULT 14
US-09-291-922-27
; Sequence 27, Application US/09291922
; Patent No. 638376
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingley, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-27

```

Query Match	2.4%	Score 65.8	DB 4	Length 1872
Best Local Similarity	44.7%	Pred. No. 1.1e-07		
Matches 302	Conservative	0	Mismatches 367	Indels 6
			Gaps	1
QY	123 GCTTCGTGTTCTTCACAGAGCCGGTGACCTCGACGATATCTTGAGGACAAGATGTCG	182		
Db	122 GCTTCGTGTCGCGCTCCCGAGCCGGGGGCGAGTCATCCAAAGAACAAAGGCAATTTCAAG	181		
QY	183 GGGGCTGTTCTTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGCAAGGGGTGGACAAT	242		
Db	182 TACGCCCTTACCGTCGCGCCCTCTGTGCTTCCATAGGCCAACCATCGTCCTCGGCTACGACGTT	241		
QY	243 GCCACCATCGCAGCTGCTGTTCTGTATATATAAGAAGAAATTCAATTGCCAAATGAGCCC	302		
Db	242 GGGGTGATGAGCGGTGCGCTCGCTGTATCATCAAGAGGACCTGCAGATCACGAGCTGC	299		
QY	303 ACTGTGAGAGGACTAATTGTGTCAATGTCACCTATTCGGCGGCCCAACATCGTTACTACATTC	362		
Db	300 ---AGCTGAGATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCTTC	355		
QY	363 TCCGGGCCATTATCAGACTCGATTGGCCGACGCCCTAAGCTTATTCTCTTCAATTCG	422		

Db 356 GCGCGGAGGACGTCGCGACTGGGTGCGCCGCGCGCTCACCGCTCGCTTCGCGGCGCCCATC 415

QY 423 TACTTCTCAGCGGCGCTCATCATGCTATGCTCTCCTAATGTCATGTCCTGCTGTGGCA 482

Db 416 TTCAACAACGGCTCCTTGCTCATGGGGCTTCGCGGTCAACTACGCCCATGCTCATGTGCGG 475

QY 483 CGCTTCGTAGATGGATTGGTAATGGCTTGGCTGTCAAGCTTGTCGCTTTGTATTTCA 542

Db 476 CGCTTCGTACCGGAATCGCGGTGGGCTACGCCATCATGTGCGGCCAGTGTACACGCC 535

QY 543 GAAATAGCCCCCTTCGGAGATTAGAGGTTTGCTGAATACACTACACAAATTCAGTGATCA 602

Db 536 GAGGTGTCCCCGCGCGCTCGGCCGCGGCTTCCTCACGCTTTCACCGAGGTTCATCAAT 595

QY 603 GGAGGAATGTTCTTGTCATACTGCATGCTGTTGGGATGTCCCTGTGCGCATCACCCGAT 662

Db 596 GTGGGCATCTCCTTGCGTACGTCTCCAATAAGCCTTCGCGGCTCCGCTCCACCTC 655

QY 663 TGGAGAATTATGCTTGGTGTGCTGCGGATACCTTCATGTTCTTGGTTTGACAATA 722

Db 656 AGCTGGCGGTCATGCTCGGCATCGGCGCGCTCCGCTCGCTTCGCGCTCATGTGT 715

QY 723 TTTTATCTTCCTGAATCTCCAAGATGGCTCGTTAGCAAAAGTGGATGGCAGAGCAAAA 782

Db 716 TTCGCATGCCGAGTCTCTCGCTGGCTCGTCATGAAGGCCGCGCTCGCGACGCCAAG 775

QY 783 AAGTGTTGCAAAAG 797

Db 776 GCCGTTCTGGCCAAG 790

```

RESULT 15
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

```

	Query Match	2.3%	Score 63.4;	DB 4;	Length 1914;	
	Best Local Similarity	45.0%;	Pred. No. 5e-07;			
	Matches 284;	Conservative 0;	Mismatches 341;	Indels 6;	Gaps 1	
QY	167	GGAGGACAAGATGTCCGGCGCTGTTCTGTGTCGCCATAGTCCGCTCCATCCGAATCTATT	226			
Db	97	GAAAGGCACGTCAAGTAGTCCCTCCATATGTGCCATCCTGGCCCTCATTGGCTTGTCAT	156			
QY	227	GCAGGGGTGGACAATGCCACCATCCAGCTGCTGTTCTGTATATAAAGAAGAAATTCA	286			
Db	157	CCTTGCTATGACATTTGGGTGTAGTAGTGAGCGGCATGTACATCAAGAAGAACCTGAA	216			
QY	287	ATTGCAAATGAGCCCACACTGTGTGAGGGACTAATTGTGTAATGTCACTTATCCGGCGCAC	346			
Db	217	TATCACGGACGTG-----AGCTGAGAGATCCTGATCGGATCCTCAGTCTTACTCGCT	270			
QY	347	CATCGTTACTACATTTCTCCGGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTAT	406			

```

Db      271 GTTCGATCCTTCGCTGGCCGCCGACGTCAGACAGATCGGCCCGCTTGACCGTCGT 330
QY      407 TCTCTCTCAATTCGTACTTCTTCAGCGGCTCATATGCTATGCTCTCTAATGCTA 466
      |||||
Db      331 GTTCGCGCTGTCACTCTTCGTGGCTGTGCTCATGAGTTTCGCCGTCAACTACGG 390
QY      467 TGTCTGCTGTGGCAGCTTCGTAGATGATTTGGTATGCTTGGCTGTACAGCTGT 526
      |||||
Db      391 CATGCTCATAGCGCGCCGCTTCGTGGCCGAGTCGATGTGGCTACGGGGCATGATCGC 450
QY      527 GCCTTTGTACATTTAGAAATAGCCCTTCGAGATTAGAGTTTGCTGAATACACTACC 586
      |||||
Db      451 GCCCGTGTACACGGCCGAGATCTGCGCTGCGGCGTCCCGTGGCTTCCTGACCACTGCC 510
QY      587 ACAATTCAAGTGCATCAGGAGATGTTCTTGTCACTACTGCATGCTGTTGGATGTCCT 646
      |||||
Db      511 GGAGGTGTTCATCAACATCGGCATCTGCTTGCTTACCTGTCCAACTTCGCGTTCGCGCG 570
QY      647 GTGCGCATCACCCGATTTGAGAAATATGCTTGTGTGCTCGGATACCTTCATTGTTCTT 706
      |||||
Db      571 CTTCCGCTCCACCTCGGCTGGCGGTGATGCTCGCAATGGCGCAGTTCCGTCGCGCT 630
QY      707 CTTTGTGTGACAATATTTATCTTCTGAATCTCCAAAGTGGCTGTAGCAAAAGTCTG 766
      |||||
Db      631 GCTCGCGCTCTGTGTCTTGTGATGCCCGAGTCGCTCGGTGGCTGTCTTGAAGGCGCG 690
QY      767 GATGGCAGAGCAAAAAGGTGTGCAAAAG 797
      |||||
Db      691 CTTGCGGACGCCAGGGCTGTGCTAGAGAAG 721

```

Search completed: January 5, 2004, 15:50:46
 Job time : 175 secs

THIS PAGE BLANK (08710)